Quantum Physics: An Overview of a Weird World

Volume II

A Guide to the 21st Century Quantum Revolution

Marco Masi

First Edition

Quantum Physics:
An Overview of a Weird World

Volume II

A Guide to the 21st Century Quantum Revolution

Marco Masi Ph.D.

First Edition 2020
Volume II

Kindle eBook edition

Note:
For the discounted black & white paperback edition:
https://www.amazon.com/dp/3948295035
Color paperback edition available: check the online stores.

Contents

Introduction

This second volume on the conceptual foundations of quantum physics (QP) is a continuation of Vol. I, which was a bottom-up introduction to the basics that, with its historical approach, focused mainly on 20th-century physics. This volume continues the journey, leading us from the 20th century and into the 21st century.

Our attempt to furnish a 'grand vision' of QP will, also this time, also address a non-academic audience that is willing to make an effort to go beyond the conventional low-level popular science portrayals, as well as a certain academic reader who is looking for the conceptual foundations of QP for which a pragmatic educational system nowadays does not allow. Again, we will try to find a middle-ground compromise between a too-sophisticated academic and mathematical approach that tends to obfuscate behind a plethora of mathematical abstract notions the deeper meaning of the concepts and physical significance of the phenomena involved and a too-simplified and naïve representation of the subject that tends to misrepresent reality as it is. The approach we adopted in the first volume was unique in the sense that it closed a void: We were neither interested in dwelling in too much technical rigor nor in telling hyped stories that the popular science outlets are eagerly seeking. We looked for the hard facts, the experiments, the empiric data that present us with reality as it is while, at the same time, we tried not to fall into the temptation to impress the reader with wild speculations. The aim was, first of all, to deliver these facts so that, once fixed, the reader is enabled to distinguish between sound scientific reasoning and pseudo-scientific blither. As far as possible, we will also maintain this approach here.

However, especially with this second part, which focuses on the modern theories of QP and particle physics and which also takes a look at cosmology, it is impossible to not mention the most recent theories of theoretical physics which, to a large extent, are not a confirmed and fixed scientific truth but, quite the contrary, are still in a development phase that could not go far beyond conjectures, hypothesis, and sometimes wild speculations. This makes this volume necessarily different in character from the previous one. Sometimes this goes so far that the question arises of whether some of these intellectual endeavors can still be considered sound scientific practice or, if a border was crossed where it is difficult to distinguish between reality and phantasy or, as Wolfgang Pauli used to say, one is talking about theories that are *'not even wrong'*. In the first volume, it was relatively easy to restrict our attention to the established scientific truths and solid empiric data that, most of the time (with the exception of the Bohr-Einstein debate and perhaps few others), did not involve personal and too-subjective preferences and could be separated by ideological influences. Meanwhile, the new-millennium theoretical physics is, in large part, plagued by uncertainties and ambiguous theoretical exercises for which scientists will almost certainly have to spend several decades determining whether they have any sound foundation. Contrary to the physics of Einstein, Bohr, Schrödinger, Dirac, or Pauli, which nowadays has been systemized into a clear and coherent descriptive frame, modern theoretical physics that goes beyond the standard model (SM) of particle physics is far from being an established science. There is no consensus on several issues and no one can pretend to have definite answers to many open questions because no established 'right' answer has been confirmed by experiments and is generally accepted by scientists in the field. However, because these are nowadays at the center of most of the modern discussions on the foundations of QP and particle physics, we included them as well. It is, therefore, also unavoidable that the author will add his own perspective, which does not always align with the mainstream perspective.

While this second part builds on the basics of the first volume (making frequent reference to it) and assumes that the reader is acquainted with the main concepts of QP, it can nevertheless be read as a self-contained work if the reader already has an understanding of the subject. Also, this time, whenever a more sophisticated proof is required, the interested and more skilled reader will find it in the appendix. Meanwhile, those who would like to skip some of the in-depth analyses given in the appendix won't lose track of the conceptual foundations necessary for further reading.

The level of complexity alternates. Some chapters are a relatively easy read and ask for almost no mathematical background (such as the chapters on the interpretations of QM, the standard model of particle physics, Bose-Einstein condensates (BEC), quantum biology (QB) and quantum cosmology). Other chapters require more effort from the self-teaching student (such as the modern experiments of quantum optics in the first section or the principles of quantum information theory).

However, unlike with the first volume, the advantage of this treatise is that each section can be considered self-contained. If readers do not feel comfortable with the somewhat-more-intense formal approach of some parts they can proceed to the next section without losing contact with the subject. Specifically, it can be said that if the first section on the which-way experiments is not of interest or is too intense, you can directly skip to section II on the interpretations of QM or section III on the standard model of particle physics. Section IV.1-4 on quantum computing can be considered independent from section IV.5-9 on classical and quantum information theory. They have been combined into one section simply because both parts deal with quantum-informational aspects. This is especially the case for section V, in which each chapter can be considered independent reading.

The fact that sections are set out in a somewhat-enumerative manner and that several chapters are a standalone part is not an editorial choice. It is precisely this that reflects the uncertain state of affairs of modern theoretical physics, which is affected by many different approaches and, to some degree, disconnected research fields. Each of these reflects the nowadays existing disparate lines of investigation which sometimes aim even at contrasting goals and/or alternating attempts to find something that still has to be found.

However, whatever kind of uncertainties reign in science, they can't stop human curiosity and our innate instinct from knowing more. It is a reason to feel it more necessary to inform, first and foremost, an audience of people who consider themselves to be auto-didacts, self-teaching students, and independent thinkers who would like to go further than what they have learned at school or university and who wonder if – and how much – truth stands behind the mass media’s sensational headlines. On the other hand, one of our main aims is also to increase awareness of the subject, which might help some of the self-educating readers avoid the kind of autodidacticism that much too frequently falls into a practice of 'crackpot science'.

Who else is this book for? For physicists who didn't focus on the foundations of QM (that is, the majority of them) but who would like to complement and refresh the knowledge they received from their dry and strictly formal college education. Indeed, most physicists could profit from this book because it elucidates several concepts and foundational and philosophical aspects of quantum theory (QT) that they almost certainly did not receive during their conventional undergraduate or graduate studies. Philosophers of science who are in control of high school math with some preliminaries in calculus and linear algebra and who have already acquired some basics of quantum mechanics (QM) could equally profit from this second volume as self-contained literature. In addition, engineers, IT students, biologists, chemists, or whatever professional category with similar technical preparation can eventually proceed directly to the reading of this work.

However, it is the author's conviction that if readers take the necessary time and are willing to make an effort to, eventually, go through some parts more than once so that the more complex concepts can sink in, they will have an understanding of several aspects of the foundations and philosophical implications of QP that, in most cases, even physicists don't have.

I. Advanced experimental tests of 'quantum ontology'

This section will deal with some of the most notorious and odd quantum optics experiments that have been performed in the last few decades and that further investigate the foundations and ontology of QP. These experiments can be considered as the continuation of Wheeler's delayed choice experiment and the interaction-free Mach Zehnder Interferometer (MZI) experiments presented in Vol. I. In principle, they could complement it as a concluding part. However, the higher level of theoretical and formal sophistication makes it more appropriate to present this information in the present volume, which addresses the advanced reader.

Apart from furnishing an overview of modern state-of-the-art quantum optics experiments, the aim of the following chapters is to further highlight the non-local aspect of quantum mechanics (QM), inviting the reader to abandon the naïve standpoint of a differentiating 'which-way' (or 'which-path') particle perspective, still frequently invoked by professional physicists, and to embrace the more holistic non-separable point of view which takes entanglement and state superposition as quantum phenomena that must be regarded seriously, not just as formal expedients. Along the way, we will also demystify the myth of temporal quantum retro-causality, according to which some experiments supposedly prove that the effect can precede the cause. While, indeed, physics does not explicitly disallow the existence of retro-causal effects, we will show that, at least so far, the delayed choice quantum eraser (DCQE) experiments that seem to suggest 'back from the future' actions can be explained without invoking alternative cause-and-effect orders other than the conventional one.

1. The double crystal experiment of Zou, Wang and Mandel

Let us begin with a quantum optics experiment that is somewhat less known to the public but that is still quite mind-boggling. It was performed by a group of physicists from the University of Rochester in 1991. We will call it the Zou, Wang, Mandl (ZWM) experiment. [1] It sets the stage as an introductory experiment which, apart from being interesting per se, will acquaint us with the ambiguities involved in a which-way ontology that imagines individual particles on deterministic paths that are supposed to be localized in space and time.

Fig. 1 shows the experimental setup. The light source is a *'single-photon source'* (or *'one-photon source'*) coming from an argon ultraviolet laser –

that is, only one photon is heading for spontaneous parametric down-conversion (SPDC) during a time interval no shorter, or eventually longer, than the time of flight through the device of the two entangled photons. Therefore, the device always contains only a couple of photons. This source sends a photon to beamsplitter BS_1, which splits it in a superposition state along two paths. One path leads (after a reflection in a mirror) the photon to a nonlinear crystal (NL_1), after which the SPDC transforms it into two entangled photons with half the wavelength of the original one, called the '*signal photon*' and '*idler photon*', labeled s_1 and i_1, respectively.

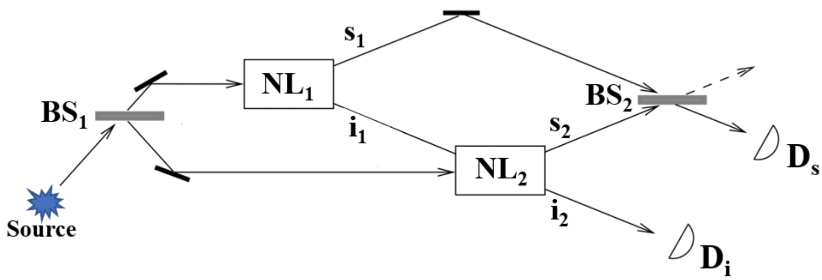


Fig. 1 The experimental setup of the ZWM experiment [1]

Without going into too many details, it may simply be said that by '*nonlinear crystals*', one indicates a more general class of optical media that can produce entangled photons (one example that we mentioned so far was the beta-barium-borate (BBO) crystals, though there exist several other types of crystals capable of producing entangled photons). They are nonlinear insofar as their physical and optical properties respond non-linearly to the intensity of the electric field of the stimulating light beam.

The other path after BS_1 leads the photon to another nonlinear crystal (NL_2), which transforms it into two entangled signal and idler photons, s_2 and i_2. One of the most important details to fix in our minds is how the idler photon i_1 of the first crystal is sent through the second crystal NL_2. In fact, if a light beam is sent at an appropriate angle, no entangled photons are produced. The crystal behaves only as a transparent medium, just like a piece of glass with low absorption. This allows for interference between the two idler photons i_1 and i_2, which can be measured at the 'idler detector' D_i. Meanwhile, on the upper stage of this optical device, the two signal photons, s_1 and s_2, are led to converge onto the second beamsplitter BS_2, where they will interfere as well. This latter interference can be measured at one side of the beamsplitter by a 'signal-detector' D_s by slightly displacing beamsplitter BS_2 from its position or inclination and changing the relative optical lengths of the two optical paths involved. (This requires very precise mechanical control on the order of less than a micrometer.) A coincidence counter (not shown in the figure) measures when detectors D_i and D_s click almost jointly. 'Almost' means that they both click during a time interval no longer than that which a photon requires to traverse the device to make certain that they, indeed, measured the signal and idler photon generated by the one and the same source photon.

In fact, in this configuration, neither detector D_s nor detector D_i can determine the path of the signal and idler photons, respectively. This is because if D_s clicks, the so-measured signal photon could have been photon s_1 coming from crystal NL_1 and transmitted through beamsplitter BS_2 or photon s_2 coming from crystal NL_2 and reflected at the same beamsplitter BS_2. Similarly, if D_i clicks, the so-measured idler photon could have been photon i_1 coming from the down-conversion at crystal NL_1 and transmitted through crystal NL_2, or photon i_2 coming directly from the down-conversion of crystal NL_2. Therefore, at first glance, it does not seem surprising that interference fringes appear, as expected, as the which-path information isn't available.

However, notice that while two entangled photons are propagating through the device, there could be only one signal photon (s_1 or s_2) and one idler photon (i_1 or i_2). Think about this carefully, and you will realize that something weird is at work. There can't exist two down-converted photons at once, one at crystal NL_1 and the other at NL_2, because that would be contrary to the fact that the source is a single-photon source. There can be only one pair of photons travelling inside the device, and these must be either photons s_1 and i_1 coming from the SPDC of crystal NL_1 or photons s_2 and i_2 coming from the SPDC of crystal NL_2. There can't be only photons s_1 and i_2 or only photons s_2 and i_1, nor could there be two signal photons (idler photons) without their idler twins (signal twins). Otherwise, that would imply that both crystals have down-converted the pump laser photon to only one signal (idler) photon. This is something that has never been observed for the crystals in isolation. Moreover, what would be entangled?

Looking at things from the particle which-way perspective, the question is: How can a signal photon interfere with another signal photon at beamsplitter BS_2 if only one of them is allowed to exist? Similarly, how can the idler photon interfere with another idler photon at crystal NL_2 if only one of them is allowed to exist? One might argue that it is like the situation in the double-slit experiment. There, also, we think of only one photon going through two slits. However, here we are speaking of the interference between two different photons which are supposed to be created by an SPDC into different places at different times, namely, one generated in crystal NL_1 and another in crystal NL_2.

The fact, however, is that one observes an oscillating interference phenomenon. Curve A of Fig. 2 shows the counting rate (per second) of photons measured at detector D_s by slightly displacing beamsplitter BS_2. A sin/cos wave clearly appears, testifying to the fact that interference indeed occurs. The interference at detector D_s is manifested independently of whether coincident signals are recorded at detector D_i. Similarly, one could show the existence of interference 'fringes' at detector D_i by slightly changing the optical path between crystals NL_1 and NL_2.

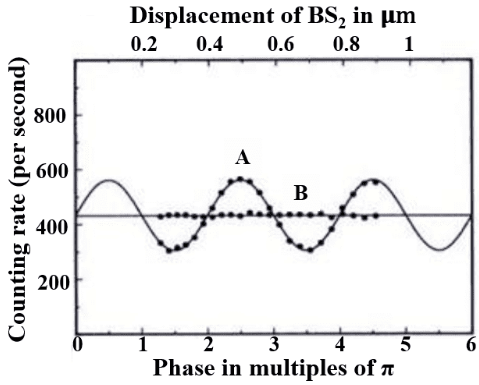


Fig. 2 Measured photon counting rate at D_s as a function of BS_2 displacement.

All this is even weirder if we keep in mind that the two processes – those corresponding to the two signal photon emissions s_1 and s_2 – are emitted at random times. SPDC is 'spontaneous' in the sense that the time of the energy transition involved at a microscopic level responsible for the creation of the entangled photons is determined by probabilistic quantum rules and there is no way to synchronize the down-conversion process of two nonlinear crystals.

To further test the behavior of this device, ZWM wondered what would happen if one interrupted the first idler path from crystal NL_1 to NL_2 – for example, by placing in between an object that fully absorbs idler photon i_1? In this case, one obvious thing is that there can be no interference at detector D_i, as idler photon i_1 which could be interfered with, is blocked physically. However, we can't say the same about the interference of the idler photon(s?) s_1 and/or(?) s_2 at beamsplitter BS_2. Blocking the idler photon i_1 should have no influence whatsoever on what happens to the signal photon and we should expect the same interference pattern emerging at detector D_s of the previous case – that is, curve A in Fig. 2. And yet, when the idler photon i_1 is blocked, the interference fringes disappear not only at detector D_i but also at detector D_s. The flat line B of measurement points in Fig. 2 appears, clearly testifying to the fact that the previous interference pattern has disappeared.

In some sense, this is not too surprising. That is because now we can distinguish the path the photons take. If detector D_i clicks and detector D_s

does not, that could be due only to the absorption of the idler photon i_2 generated by the SPDC of crystal NL_2 and coming from the lower path under beamsplitter BS_1. If detector D_s clicks and detector D_i does not, that could be due only to the absorption of the signal photon s_1 generated by the SPDC of crystal NL_1 and coming from the upper path above beamsplitter BS_1. If both detectors D_i and D_s click at the same time (that is, during the time interval allowed by the coincidence counter to rule out the possibility that other photons are propagating inside the device), the signal photon s_2 at detector D_s and the idler photon i_2 triggering D_i must have been generated by the SPDC at crystal NL_2. So, complete which-way information is present and no interference is expected, as observed.

On the other hand, if we insist on maintaining an ontology which imagines particles traveling along separate deterministic paths, this behavior is quite difficult to explain. Why does the spatial interruption between crystals NL_1 and NL_2 influence signal photons s_1 or s_2 traveling along completely different paths and making interference fringes disappear at detector D_s? We are no longer allowed to sweep the question under the carpet by saying that this is due to the fact that only one photon is arriving at BS_2 and that it could not interfere with any other photon. In the previous configuration, that without blocking the idler photon i_1, we saw how interference phenomena appear even if, according to our which-way conception, only one photon is present.

The only way out is to accept the fact that the physical object interrupting the idler photon path between crystal NL_1 and NL_2 also instantly collapses the wavefunction throughout the entire device. Don't forget that the signal and idler photons, s_1 (s_2) and i_1 (i_2), are entangled. While in the first configuration with no interruption, the collapse takes place only later, at the instant of detection of one of the two detectors D_s or D_i, in the second configuration with the interrupting object inserted along the path of idler photon i_1, the wavefunction collapses from two entangled particles, which we could naively see as being in two places at the same time, to a quantum state of two distinct and individualized particles being in two separate regions of space. At that stage of the process, there are 'really' two individualized particles, each on different paths and unable to interfere.

As we have amply discussed in Vol. I, this means they are one and the same until the wavefunction collapses. Entanglement implies indistinguishability in a more radical and fundamental sense that we should never mistake for the classical indistinguishability. Entangled – that is, indistinguishable – particles are one and the same object until observed and can't be described by the sum of subsystems that one would like to analyze separately.

Moreover, recognize how not only entanglement is at work but quantum superposition as well. What we must conceive of being 'superimposed' is the SPDC of the two crystals. We know that particles can be in a superposition of quantum states, such as the spin-up AND spin-down states of the electron. Here, the down-conversion process in both crystals is 'on' AND 'off' at the same time. (Of course, that happens at a microphysical sub-atomic level, not for the entire crystal.) As long as the time of triggering one or both detectors hasn't come, there is a state of superposition of the two entangled photons emitted from crystal NL_1 and the two entangled photons emitted from crystal NL_2. However, once a detector collapses the state function, it is not possible to detect a photon from each crystal and, finally, only two (not four) photons will be measured. By the way, this also shows that the physical process of SPDC does not collapse the source photon that traverses the crystals and does not spatially collapse the wavefunction but, rather, transforms it from the state of a single photon to that of an entangled photon.

In some of the following experiments, we will see additional examples of the coexistence of quantum entanglement and superposition occurring at the same time. The point is, we must take these seriously, not just as an abstract representation of facts. As long as instantaneous state reduction does not come into play, only one object is propagating along all paths, namely, that from a point of emission in the source to one of the detectors. In between, and during, the emission-detection time interval, we can describe the 'state of being' of this entity only with an abstract state vector, though any reasoning based on a counterfactual definiteness pointing at some separate object with definite properties in a dividing and separative space-time conception is misplaced. In a certain sense, speaking of 'entangled particles' is a self-contradiction in terms, because we still picture, in our minds, two separated particles somehow interweaved throughout space, though one can't do otherwise due to the limitedness of human language. A more appropriate understanding might furnish us Feynman's path integrals approach (see path integrals and Feynman's diagrams in Vol. I), in which one considers the final trajectory of a particle as resulting from the interference of all the possible paths, that is, a sum over 'histories'. Feynman's calculation technique determines the probability of a particle traveling from a space-time point to another (the propagator) and assumes that it travels along all the possible paths allowed at once. It does not work with anything being a corpuscle; it simply adds all the wavefunctions describing the possible histories that we, by counterfactual definiteness, imagine to be the paths of a single particle.

What we should evince from this paradigmatic experiment is that the most sensible way to interpret the facts is to give up a particle model and embrace a more integral and holistic perspective. What spans the experimental set-up of this quantum optics experiment is never a particle, and not even a wave, but a wavefunction, a probability wave, a not-better-defined 'physical entity' that takes all the possible paths through the device at once and, at the instant of measurement or absorption, collapses. Or, to put it in other words, if you prefer, one might complement this view by that which is contrary, namely, by conceiving of this entity as taking neither 'this path or that path', nor that of traveling 'this path and that path', because there is only ONE path.

At any rate, the experiment of ZWM should instill some doubt, to say the least, as to whether it still makes sense to insist on a which-way interpretation of QM, where one imagines corpuscles (photons or material particles) possessing definite properties, localized in space and time and traveling along well-defined deterministic paths, as our classical Newtonian mindset would like to believe. If, instead, we give up this model and begin to realize that there are neither particles nor waves, this might lead us a step further.

In this chapter, we considered ZWM's experiment as an 'appetizer' that paves the way to other experiments which will further highlight the subtleties and deep conceptual implications of modern QT. The following chapter will shed more light on the extent to which the which-path perspective is appropriate and, at best, only complementary to a more integrating and non-dual perspective.

2. The which-way reconsidered.

Even among several physicists and philosophers of science, it is considered a common wisdom that whenever an experiment is performed that attempts to gain insight into a particle's path, that is, its 'which-way' information, then automatically all interference phenomena must disappear. While this interpretation isn't incorrect, it carries with itself some ambiguities that are potentially misleading, especially when it comes to the celebrated DCQE experiments that the next chapters will illustrate. In this chapter, we will clarify the deeper meaning of the 'which-way' (or 'which-path') rule and how far we are allowed to effectively think in terms of deterministic paths resorting to local realism. This will prepare us to tackle the supposed effects of temporal quantum retro-causality which are sometimes invoked incorrectly. There are several methods, interpretations, and approaches that one can use to explain the DCQE experiments in a more appropriate context. Here, we will adopt that taken by David Ellerman [2]. (Other, equivalent approaches that demystify retro-causality exist as well [3] [4], [5].)

Let us go back to the very basics, which told us how the interference fringes of the classical double-slit experiment come into being. (Recall Young's double-slit experiment in Vol. I.) Please keep in mind how that formulation, which is also what one finds in most textbooks, holds only when the two interfering beams from the slits have the same polarization – that is, when the two field vectors are co-directional on the polarization plane at every instant. If one studies the conventional Young double-slit experiment, these are only subtleties that one can ignore, as the two slit-beams emerge from the very same incident beam and, therefore, always have the same path difference and polarization. However, in double-slit experiments involving polarizers, these aspects can no longer be neglected. Moreover, by inserting polarizer filters in front of one or the other slits, one must also consider the extra phase shift that these could eventually induce on the path of the respective slit. If these are taken into account, the considerations involving interference patterns require further attention and its form requires a slightly more complex representation, which, however, will allow us to gain much deeper insight into the meaning and correct interpretation of the 'which-way' experiments from the quantum mechanical perspective.

Let us illustrate this in detail beginning with Fig. 3. Suppose the light coming from a light source is polarized along the horizontal direction (0° polarization, by convention) with a polarizer in front of both slits. Then, if one were to place a linear polarizer after one of the two slits (say, S_1), this would change the polarization state of the beam from S_1 relative to that of the beam coming from slit S_2 (still in the 0° polarization state) by an angle $\delta\theta$. (The symbol δ will always imply a relative difference between two quantities – here, the relative angular direction difference of the polarization vector between the two beams.)

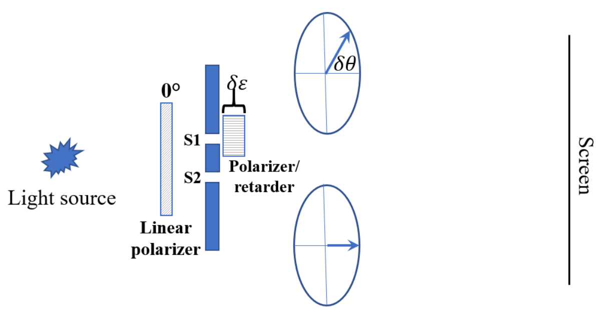


Fig. 3 Polarization angle difference $\delta\theta$ and phase shift $\delta\varepsilon$ between light beams coming from slits S_1 and S_2.

Moreover, a polarizer, as any transparent plate with some optical refractive index (that is, light is slowed down), will also cause a change in the optical path – here, a retarding relative path-phase shift of the beam from slit S_1 relative to that of slit S_2 by an amount $\delta\epsilon$.

The latter quantity is usually expressed in degrees (or radians) as a phase shift proportional to the wavelength. For example, a half wavelength $\frac{\lambda}{2}$-shift is identical to a $\delta\varepsilon = 180°$ (or $\delta\varepsilon = \pi$ radians) shift. It can be shown (see Appendix A II) that this change in phase and polarization will also lead to a different double silt interference pattern than that to which we were accustomed in the conventional Young experiment and that can be captured by modifying the intensity function of the double-slit beams interfering on the screen (as a reminder, see the Appendix of Vol. I on trigonometric functions and waves and complex numbers) as follows (for the sake of simplicity, the dependence on the x-coordinate is omitted):

$$\begin{aligned} I(\delta\phi, \delta\varepsilon, \delta\theta) &= I_1 + I_2 + 2\sqrt{I_1 \cdot I_2} \cdot cos(\delta\phi + \delta\varepsilon) \cdot cos(\delta\theta) \\ &= 2I_0[1 + cos(\delta\phi + \delta\varepsilon) \cdot cos(\delta\theta)]\,, \qquad \textit{Eq. 1} \end{aligned}$$

where, as usual, $\delta\phi$ is the angular phase path difference of the beams that determines the angular dependence along the vertical screen direction while the last passage simplifies the expression assuming $I_1 = I_2 = I_0$, that is, with the two identical slits transmitting the same amount of light with intensity I_0. Strictly speaking, the real intensity behavior we show in the graphs is obtained by multiplying Eq. 1 with an exponential damping factor (such as, for example, $e^{(\delta\phi/10)^2}$). As we know (see the many slits interference and diffraction in the appendix of Vol. I), the realistic intensity function of the fringes is damped out by an angular term, the diffraction envelope. We will maintain this in the graphs but, to keep things simple, will not consider it in the text equations because it isn't relevant to the aspects we will point out.

Let us inspect Eq. 1 and understand how it works. Note that both cosine functions still exclusively modulate the interference term. If one or both of these is zero (say, for example, for $\delta\theta = 90°$), only the first two terms are left ($I = I_1 + I_2 = 2I_0$). This aligns with the case in which the two slits' intensities are added without interference, corresponding to the normal distribution curve, as we shall see next. Meanwhile, when one of the cosine functions equals -1 and the other +1, for there is full intensity subtraction at some point on the screen: $I = 0$, it is a 'no fringe' minimum. When both are +1 or -1, there is full intensity summation: $I = 4I_0$, the fringe peak. If you wonder how it could be that two slits can produce an intensity peak of four slits, just consider how energy must be conserved. The energy is not lost but only shifted from the Gaussian distribution minima locations, concentrating it due to interference on the peaks and, so to speak, 'piling' it up on the maxima.

To visualize this, we can play around with phase shifts and polarizations for different experimental double-slit setups that involve polarizers.

If we change the path phase difference at one of the slits ($\delta\varepsilon$ in Eq. 1 varies) by introducing a 'retarder' plate, for example, a polarizer that does however leave the polarization of the beam unaffected ($\delta\theta = 0$), or just several retarding transparent media with different thicknesses, then one obtains the interference patterns as shown in Fig. 4.

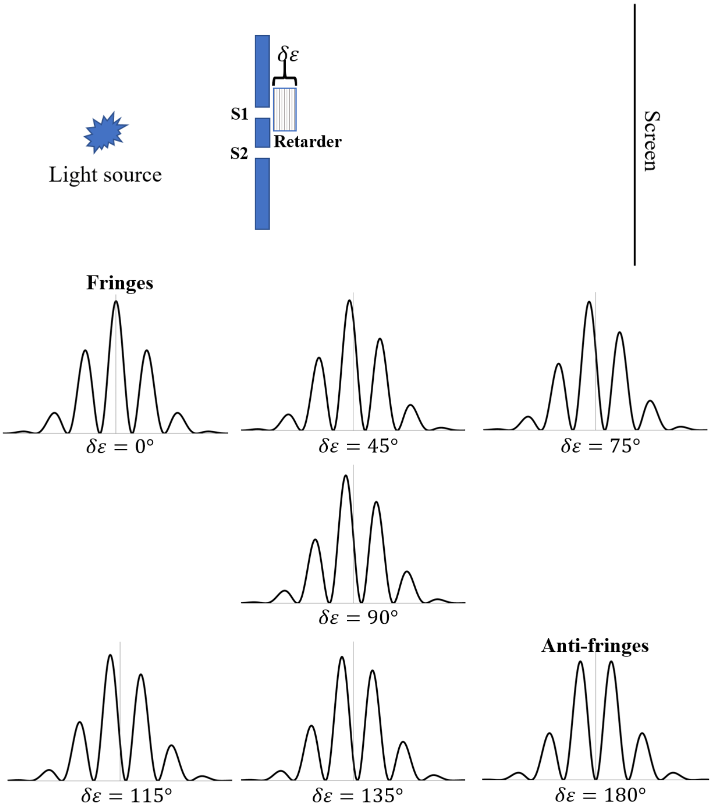


Fig. 4 How the double slit interference pattern changes due to a phase shift difference $\delta\varepsilon$ *imparted by a 'retarder' plate.*

The first interference pattern ($\delta\varepsilon = 0$) represents the conventional double-slit fringes interference with no retardation and no polarization difference. Varying the phase shift between the two slits causes the interference fringes to shift as well. The central most intense fringe is no longer centered with the geometric horizontal axis of the two slits.

The most interesting situation for us is that of a phase shift of 180° (half a wavelength path phase shift). In such a case, in opposition to the conventional fringes, the maxima become minima and the minima become maxima – that is, the white fringes are replaced by the black ones and vice-versa. These fringes are therefore called '*anti-fringes*'.

Compare the above result, obtained exclusively from the perspective of an optical electromagnetic (EM) wave, to the which-way perspective. As long as the polarization of the two beams is the same, and whatever phase shift one applies to the wave/photon traveling through the slits, the interference never disappears. This aligns with the idea that we can't gain any information about which path a photon takes only from the phase. The phase difference can't be used as a marking method for one single photon going through one or the other slit. The concept of the phase is a relative, not an absolute, one. It makes no sense to speak of a phase of one photon or a wave. Rather, only a phase difference between two waves is a meaningful physical quantity.

Note, however, one interesting and decisive fact: If the first fringe pattern is added to the last anti-fringe pattern of Fig. 4, the resulting curve is the usual Gaussian bell-shaped normal distribution (Fig. 5), which we know to be the curve that appears when we attempt to obtain the particle's which-way information or the diffraction pattern which appears for the 'single-slit' or the pinhole with an aperture size comparable to the wavelength of the incident light. (See the chapter on Heisenberg's uncertainty principle in Vol. I.)

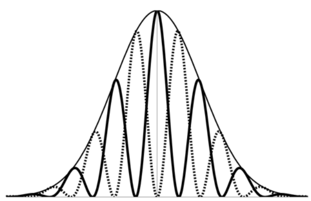


Fig. 5 The normal distribution as a sum of fringes and anti-fringes.

In all the previous experiments and considerations, we have always taken the disappearance of the interference fringes into the bell-shaped curve as a sure sign that we have switched from a wave-like behavior to a particle-like behavior. And, according to the which-way interpretation, the existence of fringes signals the wave-like character of a particle but does not allow for any information about which path it went along (i.e., through which slits or through which arm of an MZI). Meanwhile, the clumpy curve, which is the diffraction envelope itself, signals that a logical inference about the path is allowed, but the interference is inevitably lost.

Here, however, things look much more subtle. It turns out that the pattern that lacks the interference fringes, which manifests due to a lack of knowledge of the path the particle takes, hides the information by overlapping the complementary fringe- and anti-fringe interference patterns that appear when this knowledge is available. We might say that the particle behavior resulting from the manifestation of the diffraction envelope is a combination of wave-like phenomena in disguise. Is this simply a mathematical coincidence or does it have a deeper meaning?

Let us analyze the opposite case, that with a difference in the polarization of light between the two slits ($\delta\theta$ in Eq. 1 varies) but no phase shift ($\delta\varepsilon = 0$). One possible way to obtain this could be, for example, by inserting before the slits a linear polarizer and in front of both slits two identical '*half-wavelength plates*' (HWP). HWPs also shift (retard) the path phase by $\frac{1}{2}\lambda$ but their main function is to rotate the linearized light by twice the angle between the fast axis and the polarization vector. Waveplates are characterized by a '*fast axis*' and a '*slow axis*' along which the polarization component travels faster or slower, respectively. For example, if a polarized light beam enters an HWP with its polarization vector tilted by 45° relative to the fast axis, the outcoming beam will be tilted by 90° – that is, it will transform from 45° to -45° polarized light (see Fig. 6).

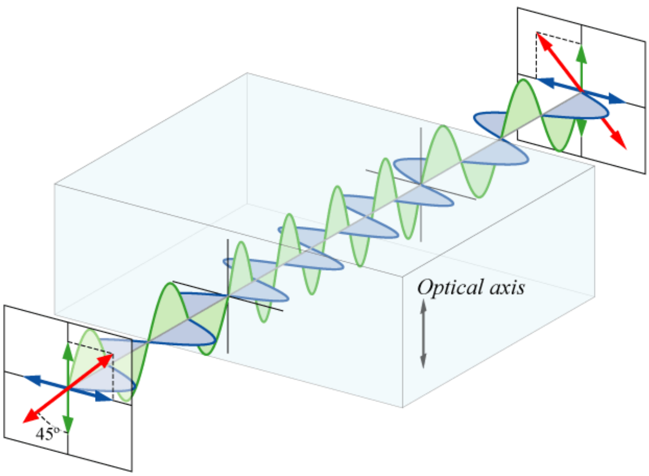


Fig. 6 90° Rotation of the polarization vector with a half-wave-plate.

Please keep in mind that, from a quantum physical perspective, polarizers and wave plates are ***not*** measuring devices. We should not associate an operator with them, at least not in the sense of an observable that causes state reduction. They do not 'measure', 'collapse', or 'reduce' anything. They 'select', 'convert', or 'change' the evolving quantum state of a particle and its related state function before a measurement takes place. This is an important distinction to which we will need to pay further attention later.

Let us apply the HWP for our purposes here. The linear polarizer's function is to fix one polarization vector of the incoming light, say, at 0° relative to the fast axis of the HWP. Two HWPs are needed in front of each slit to compensate for the retarding optical path phase shift leading to a net phase difference of $\delta\varepsilon = 0$. One HWP is left fixed at 0° while a physical rotation of the second HWP by an angle α causes a polarization angle difference between the first and second slits of $\delta\theta = 2\alpha$. The resulting different interference patterns are shown in Fig. 7.

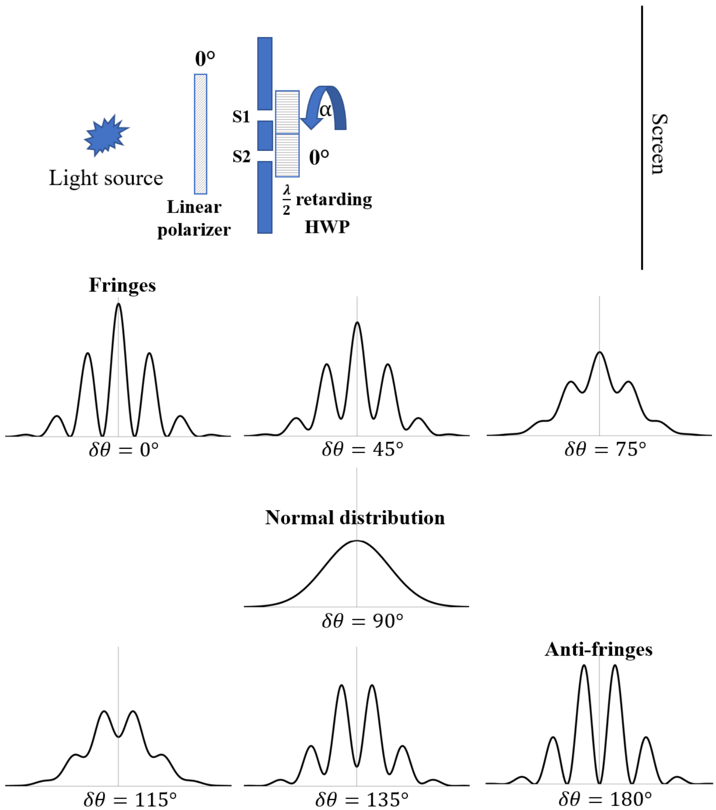


Fig. 7 How the double slit interference pattern changes due to a polarization shift difference $\delta\theta$ imparted by rotating the WHP.

One can observe a different behavior than the phase-shift fringes patterns of Fig. 4. The peaks are not shifted due to polarization but their amplitude is modulated. Again, $\delta\theta$ =0 is the conventional double slit situation (the same as $\delta\varepsilon$=0 of Fig. 4). For the orthogonal polarization ($\delta\theta$ =90°), one obtains the normal distribution density probability function, which was not (explicitly) present in the phase-shift case.

The normal distribution resulting for the orthogonal polarization is half the height of the fringes' central peak. This is obvious if you consider that the area under the interference curves (the integral, that is, the total photon count) for either the fringes, anti-fringes, or normal distribution case must always be the same for each due to energy conservation.

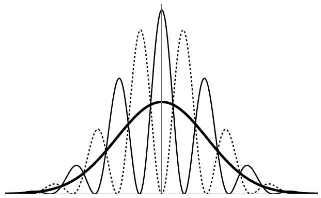


Fig. 8 Fringes and anti-fringes overlay, and the normal distribution for $\delta\theta$ =90° of Fig. 7.

This time, a special situation arises in which the interference disappears for the orthogonal polarization. This aligns with the which-way perspective. In fact, in contrast to phase shifting, from what we have seen in the which-way experiments of Vol. I, we well know that polarization can be used as a means of 'marking' a photon to gain information about which slit (or MZI arm) it went through and, with this, to resort to retro-ductive reasoning about the path we imagine it has taken. Again, the wave-perspective aligns one to one with the corpuscular which-way perspective.

To investigate this equivalence further, let us consider the special case illustrated in Fig. 9.

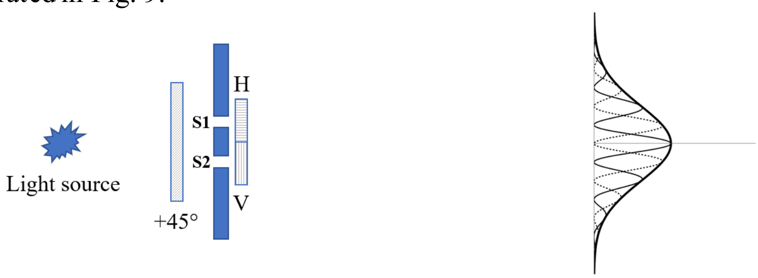


Fig. 9 Double slit with photon-marking horizontal and vertical polarizers.

A linear polarizer with diagonal 45° polarization (say, relative to the horizontal x-axis) is inserted before the two slits, while a horizontal polarizer and a vertical polarizer are inserted in front of slits S_1 and S_2, respectively. From the wave perspective, this is simply another example of the aforementioned orthogonal polarization case ($\delta\varepsilon = 0;\ \delta\theta = 90°$). From the which-way perspective, because photons are therefore marked, the interference fringes disappear and the bell-shaped curve forms on the screen. According to the quantum mechanical formalism, the state of a single photon going through this arrangement of polarizers and slits must be in a superposition state of a photon going through slit S_1 with horizontal polarization H and slit S_2 with vertical polarization, namely:

$$|\Psi\rangle = \frac{|H\rangle_{S_1} + |V\rangle_{S_2}}{\sqrt{2}}. \qquad \textit{Eq. 2}$$

The vital detail we must always keep in mind is how the which-way perspective must be interpreted correctly. One could take two approaches.

The first one thinks of the particle as going through both slits and interfering with itself only as long as we do not try to find the slit it went through, that is, its which-way. This which-way conception, however, insists on the particle-like idea and implicitly suggests that what the polarizers do is reveal to us through which slit the photon went once we measure its polarization. It seems so obvious to us that, if the outcoming photon has a horizontal (vertical) polarization, it must have gone through the first (second) slit. That is, one imagines, again by a retro-ductive cognitive act of counterfactual definiteness, that the wavefunction may have been collapsed into a particle state earlier, at the stage in which it went through the +45° polarizer, and that then the photon has gone through one – and only one – slit, finally manifesting this information to us in case we measure its polarization with the H/V polarizers. Even if we would not have placed the polarizers after the slits, by a mental projection we nevertheless imagine the photon going through one or the other slit, though we would then never know which.

The second interpretation of this state of affairs, which is less prone to accepting such a naive quantum ontology, thinks of the same particle going through both slits. It doesn't forget that, according to QP, **slits *do not separate*** a quantum object along different paths but superimpose its quantum state and that **polarizers are *not* measurement devices** that project or collapse but instead only select or change the evolving state vector. As long as there is no measurement, that is, the collapse of the state function, due to an interaction with a sensitive measurement device that allows for a readout, the quantum system is still in the state described by the wavefunction (or state vector). In fact, a polarizer reveals nothing (no operator, no observable is acting on the wavefunction) unless we do not place in front of it a detector (such as a photomultiplier, a CCD camera, a photodiode, etc.) which absorbs the photon and by which, then and only then, state reduction as signaled by a 'click' or a readout of a value becomes possible (the eigenvalue). Before that instant of the collapse, we can't conceive of any particles flying separately along a path with separate polarizations. Rather, we must still think of the wavefunction describing the system as being in a state of real superposition as a whole inseparable and unique entity being both here and there and both having one and the other polarization.

It is easy to show that the second conception must be taken seriously. Simply place, in front of the experiment of Fig. 9, another diagonal 45° polarizer, as shown in Fig. 10 top. Interference fringes reappear. When the quantum erasing polarizer is rotated by 90°, that is, by orienting it along the -45° direction, as in Fig. 10 bottom, interference is still present, but the anti-fringes will appear instead (for any polarization angle between these perpendicular directions, the intermediate cases of Fig. 7 appear, something we won't dwell on any longer here).

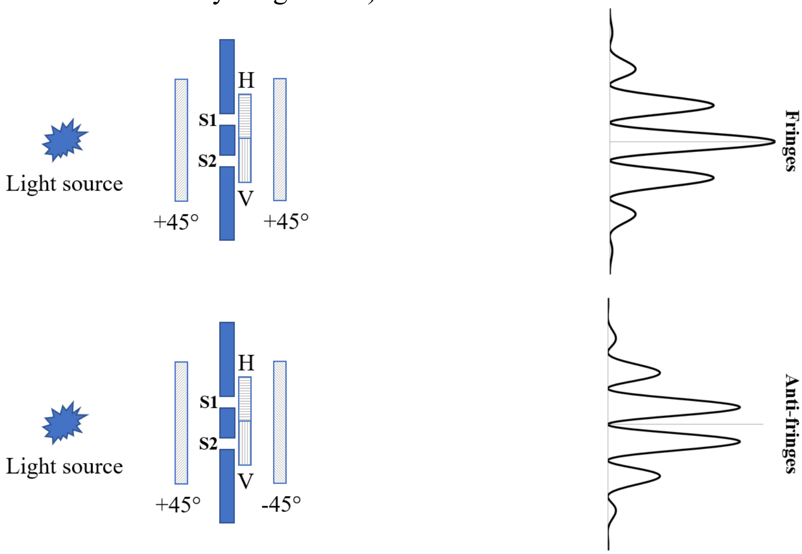


Fig. 10 A diagonal polarizer acting as quantum eraser in the experiment of Fig. 9.

From the wave optics perspective, this is a quite obvious fact because the difference in polarization is cancelled ($\delta\theta = 0°$): All photons will emerge with the same $\pm$ 45° polarization. Also, from the which-way perspective, everything looks fine because the photon's marking has been lifted – that is, the last polarizer acts as a quantum information eraser that makes the two slits' paths indistinguishable; the fringes reappear, as expected. (Recall that we had a similar situation with the reappearance of the spin superposition states in the MSG experiment; see the chapter in Vol. I where we questioned whether information is fundamental.)

This should give us food for thought with regards to the above two interpretations. First, note that if there were truly a photon taking a definite path through one or the other slit, then, when it traveled farther until it encountered the $\pm$45° polarizer acting as a quantum eraser, it must, from that point on, 'forget' where it came from, and the quantum eraser must have acted retro-causally in the past!

This is because, if the first conception which imagines particles going through one or the other slit with a definite polarization caused by a state projection of the two orthogonal polarizers were correct, there could be no recovery of the fringe or anti-fringe pattern later. If a photon truly is going through a slit with some polarization 'out there', this would have been the end of the story. The second diagonal polarizer cannot recover any interference pattern or reproduce it out of the blue.

For example, say the photon went through the slit with the horizontal polarizer; then, in the short time interval during which it traveled from the H polarizer to the ±45° polarizer, we would conceive that it had gone through one and only one slit. It 'says': *"I went through only slit S_1 and will not be able to interfere with a copy of myself coming from slit S_2. I will have to hit the screen according to a normal distribution"*. However, shortly after, it encounters the ±45° polarizer, and this information about which slit it went through will be erased. At that point, how can the single photon 'change its mind' and distribute itself on the detection screen according to a fringe (or anti-fringe) probability wave if it has no copy of itself from the other slit with which to interfere? This seemingly paradoxical state of affairs forces those who do not abdicate from a particle conception to assume that there must be some sort of temporal quantum retro-causal effect according to which the particle that encounters the quantum eraser in the present sends some information back into the past to itself, before it traversed the slit, 'telling' it to go through both slits instead of only one, in order to recover the interference pattern.

Retro-causality is not forbidden, in principle, according to the current known laws of physics. However, this should, at a minimum, lead us to some eyebrow-raising. And, for those who abhor nondeterministic interpretations of QM without hidden variables, not all hope is lost: It is possible to cerebrate elaborate models such as the 'De Broglie-Bohm pilot wave theory' (also called 'Bohmian mechanics' (BM)), which could, in principle, save the appearances and explain all this without quirky 'backward-in-time influences' that restore interference fringes, and, nevertheless, maintain a deterministic particle-like ontology. We will take a look at this and other interpretations of QM in the dedicated chapter later.

However, the question is: Is it really necessary to resort to retro-causation and/or preserve determinism to explain the observed facts? The answer is simple: It is not at all necessary if we give up the idea of an ontology describing point-like particles traveling definite paths and being localized in space and time.

Such an idea arises due to our unaware misunderstanding of how things work, which we might call the '*separation and measurement fallacy*' ('separation fallacy' being terminology suggested by David Ellerman [2]). In this fallacy, we imagine something individualized and separated into two or more paths and/or measured when it is not. We must simply accept that the two slits do not 'separate' anything; they only create a superposition of states. Otherwise, this would presuppose the collapse of the wavefunction to an eigenstate, whereas, at this stage, there is still none. The same applies to the polarizers: They do not collapse the wavefunction and they are not measurement devices. One measures with a detector that provides a readable eigenvalue, which a polarizer does not do. If we conceive of a wavefunction as describing the system as a whole, not separable into subsystems, describing a physical entity propagating towards the screen in a state of 'real' superposition (that is, 'real' in the sense that Eq. 2 expresses an ontology, not just a state of ignorance), then the paradox dissolves naturally.

Moreover, note the analogy with the experiment of ZWM in the previous chapter I.1. On that occasion, we also dealt with a physical situation in which a single photon stream at a beamsplitter (the signal photons s_1 or s_2 at beamsplitter BS_2 of Fig. 1) nevertheless produced a wavy interference pattern. We wondered how one particle could interfere if it did not have another particle with which to interfere. We showed that if we wanted signal photon s_1 to interfere with a second signal photon s_2 at beamsplitter BS_2, then the SPDC of both nonlinear crystals NL_1 and NL_2 must come into play. However, this is impossible because only one source photon at a time is produced, which cannot lead to four down-converted photons (two signal and two idler photons) due to simple considerations of energy conservation. Retro-causality didn't even enter our minds because it wouldn't explain anything. The simplest and most natural conclusion was to give up the separative space-time conception.

In conclusion to this chapter, it is instructive to see how this is also manifested in the formal description of QM, of which we will take extensive advantage in the coming discussions. Recall how we defined the vertical and horizontal polarization vectors (see the chapter on the quantum superposition principle of Vol. I):

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}}\,|\nearrow\rangle + \frac{1}{\sqrt{2}}\,|\searrow\rangle,$$
$$|\uparrow\rangle = \frac{1}{\sqrt{2}}\,|\nearrow\rangle + \frac{1}{\sqrt{2}}\,|\nwarrow\rangle\,.$$

Notice how, also, other equivalent geometrical representations are possible. For example, because vectors can be represented equivalently by their reflections, one can equate $|\nwarrow\rangle = -|\searrow\rangle$ and we can rewrite the vertical polarization as:

$$|\uparrow\rangle = \frac{1}{\sqrt{2}}\,|\nearrow\rangle - \frac{1}{\sqrt{2}}\,|\searrow\rangle.$$

Adopting this alternative representation, one can highlight the anti-symmetric components. Let us use degrees instead of arrows and define the angles relative to the horizontal x-reference axis. Then we can rewrite the two polarization vectors as:

$$|H\rangle = \frac{1}{\sqrt{2}}|45°\rangle + \frac{1}{\sqrt{2}}|-45°\rangle, \quad \textit{Eq. 3}$$

$$|V\rangle = \frac{1}{\sqrt{2}}|45°\rangle - \frac{1}{\sqrt{2}}|-45°\rangle. \quad \textit{Eq. 4}$$

Expressing the same quantum state of the system in the 45° diagonal basis, that is, inserting these into Eq. 2, it 'splits' into the sum of two terms:

$$|\Psi\rangle = \frac{|45°\rangle_{S_1} + |45°\rangle_{S_2}}{2} + \frac{|-45°\rangle_{S_1} - |-45°\rangle_{S_2}}{2} \quad \textit{Eq. 5}$$

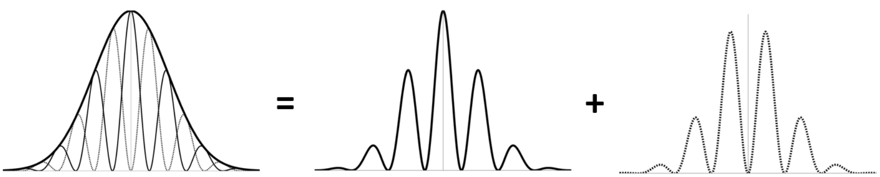


Fig. 11 Graphical representation of Eq. 5.

The first right-hand side term of Eq. *5* represents the symmetric wavefunction while the second one represents the anti-symmetric. We know now the significance of that negative signature of the second term as indicating an anti-symmetric wavefunction (see also Bosons, Fermions, and Pauli's exclusion principle in Vol. I). The $\frac{1}{2}$ coefficients tell us that there is a 25% probability of obtaining one or the other outcome instead of 50%. This is because a diagonal polarizer filtering horizontal or vertical polarized photons will block 50% of them.

What we have done is create a change of the eigenvectors basis from the H/V to the 45°/-45° basis. (There are much more rigorous and formally precise quantum algebraic methods with matrix and group representations by which to do this than what we have tentatively done here, but this should not come as entirely new information; it is, in essence, the same operation we completed in Vol. I in the chapters on spinors or the superposition principle.) One represents the very same quantum state in a different eigenbasis – or, to put it in a more intuitive language, we are 'looking' at the state of the quantum system no longer along the horizontal and vertical directions but along the two diagonal ones. However, the physical state of the system remains unaltered. By doing so, one discovers that the quantum state of a photon emerging from the polarization plates in the setup of Fig. 9, which led to the normal distribution diffraction pattern, is a quantum superposition state of the fringe and anti-fringe interference patterns. What the insertion of the 45° or -45° diagonal quantum erasing polarizers do in Fig. 10 is 'select' and 'filter out' from the bell-shaped distribution the anti-fringe or fringe, respectively. They do not 'collapse' anything. This, obviously, recovers the interference pattern characteristic of a wave.

So, finally, we can summarize the last two chapters as follows. The separation and measurement fallacy rests on the (more or less unaware) assumption that state reduction occurs earlier than the measurement of a device and that the detector reveals only what was already present. Instead, what all this must tell us is that the photons that make it through in Fig. 9 and that are 'marked' with the H/V polarizers, are set into a $\pm 45°$ superposition state only along the diagonal polarization orientations where the single photon is still, so to speak, a 'fringe-photon' and 'anti-fringe photon' at the same time. The polarizers modify the system's quantum state function while it is evolving towards a measurement device, but they do nothing that can be compared to a measurement on the incident beam. They don't even provide any information, as this is something that arises only at the time of the act of measurement – that is, when state reduction occurs. Prior to that detection, nothing exists in one or the other eigenstate. Projection or collapse happens only at the very end of the chain, when the detector 'clicks'. The $\pm 45°$ superposition evolves until it hits a detector and some distinction is made – that is, it selects the fringe or anti-fringe state and then the single photon hits correspondingly the fringe or anti-fringe. This is a distinction we call 'measurement' or 'detection'. A measurement can be defined as an 'act of distinction'. In a certain sense, we might even say that quantum erasers do not 'erase' anything. They only 'change', 'filter', or 'select'. The so-called 'which-way' information cannot be 'erased' because no particles were traveling along one or another path in the first place. And even less do they lead to any retro-causal actions into the past. After all, speaking of 'which-way' experiments and information 'erasing' devices is bad terminology.

3. The complementary principle and the Scully-Englert-Walther quantum eraser

It is time to look in more detail at the so-called '*complementarity principle*' of quantum mechanics. Let us also use this as a chance to recollect some of the facts we have learned so far, in this volume as well as in the former one.

Complementarity refers to the fact that QM is contextual. As discussed previously (and also highlighted with the MZI experiments in Vol. I), the result of an experiment that tests specific properties of a particle, or a quantum system, depends on the context – that is, the arrangement of the experimental setup. Therefore, in QM, only the whole set of possible arrangements and observations will form a complete description of the quantum object under measurement. Each aspect is not exclusive but is complementary to the others. Bohr called it the '*principle of complementarity*'.

The typical example we know well is that, according to a specific arrangement, you will reveal the wave nature of a quantum system, whereas in a different experimental context, it may behave as a particle, but you cannot see both at the same time.

The wave-particle duality is not the only example. Think of the spin property of particles and the SG experiment. We saw that the spin along one axis does not commute with that along another spin axis, which means you cannot measure, at the same time, the spin along the x-axis and the spin along the y- or z-axes. Only one spin component can be measured, leaving the others completely undetermined. Once again, it is the experimental context that determines, by measurement, in which eigenstate the system will be projected, leaving the other observable in state superposition. In this sense, the spin components of a particle are 'complementary' to each other, just as the wave-particle aspect can't appear at once – and nor can the two spins be definite at the same time.

Heisenberg's uncertainty principle is another important example. In the same way, you can't determine the momentum and position of a particle at the same time, in what is expressed formally by the non-commutation relation of the space and momentum operator observables. In this sense, precise measurements of the position and momentum aren't possible because these are 'complementary' properties of a particle.

The question, however, is: Can complementarity be explained away by the uncertainty principle itself? Here, we are again confronted by a problem that is similar to – if not the same as – the problem we already analyzed. We saw that the appearances of the wave or particle nature of photons cannot be ascribed to the interaction or perturbation of the measurement system with the measured object. And we could ask ourselves again whether the fact that we cannot determine, at the same time, the intrinsic angular momentum of a particle might be due to a physical interaction and perturbation caused by the measurement. Could it be the case that the attempt to measure the spin of a particle causes, somewhere and somehow, a small perturbation that flips its spin value along another axis? If so, the entire principle of complementarity would rest on Heisenberg's uncertainty principle; that is, the uncertainty principle would be more fundamental than the complementarity principle, as the latter would be simply a consequence of the former. This was, and still is, a belief also held (more or less implicitly and subconsciously) by many physicists who are not trained in the foundations of QP.

However, with the SG experiments discussed in Vol. I, we showed that this is now a difficult – if not impossible – conjecture to defend, and that it does not stand up to the proof of facts. We saw that, with the application of the MSG apparatus, which is essentially a quantum eraser system, it is possible to 'restore' the spin state along an axis if we build the experimental measurement set-up, that is, if we frame a particular experimental context that does not allow for a which-way information retrieval or, more precisely, the evolving state function is maintained in superposition, avoiding the existence of separate particle paths in the first place. We concluded that this implies that we cannot think of the spin commutation relations being a consequence of interaction. There is no interaction or perturbation of the H/V filters on the photons emerging from the two slits that could explain it. We reached the same conclusion with the interaction-free which-way and Wheeler's delayed choice experiments: The way in which Nature manifests the properties of a quantum system depends on how we ask the question, not on the fact that we weren't gentle enough and supposedly perturbed the system by interacting with it. We found this again in its photonic version in the previous section. The addition of the $\pm 45°$ polarizers in Fig. 10 restored the interference patterns.

However, most of the experiments realized in practice were conducted with photons or particles which are believed to be elementary, like electrons. One might legitimately suspect that all these strange quantum paradoxes arise due to the fact that photons – that is, objects we imagine to be sort of evanescent light waves, or light particles with zero mass – might possibly have some ghostly property whereby they can act non-locally and have sufficient 'plasticity' to transform themselves, displaying first a particle-behaviour and then a wave-behaviour, and apparently even looking into the future to see what the experimenter's choice will be. However, we can hardly imagine that to be a property of a composed material object like an atom or an even bigger material object, can we?

At this point, then, we might wonder whether systems composed of material particles like atoms would display different behavior. Well, you should already know the answer. This is because, according to the de Broglie relation $\lambda = \frac{h}{p}$, there is a correspondence between the wavelength λ and the momentum p of anything with mass, like atoms, or even a large molecule; there is no physical or logical restriction to consider the momentum p of a composite and large material object. We know that electrons – particles with mass – could be diffracted by a lattice, giving rise to the Bragg-refraction and the corresponding interference pattern. We also mentioned how this has been done with macromolecules as large as 2000 atoms. [6] We should therefore not expect that so-called 'material' objects will be likely to act differently than photons. However, on the other hand, can we imagine a chunk of matter as tiny as an atom to be diffracted, going through two or many slits at the same time, and apparently changing from a particle to wave-like behavior shortly before hitting a detector or a screen, according to our delayed choice? As if this wasn't difficult enough to grasp with photons, it all enters even more conflict with our intuitive notion of what matter is.

So, to be sure that we have it right, we now analyze an experiment performed in 1991 by Scully, Englert, and Walther, subsequently known as the SEW experiment. This experiment is a synthesis of several experiments we have seen so far, but it goes beyond them; it combines, in a fascinating manner, the wave-particle duality, the delayed choice experiment, and quantum erasing for atoms. It uses atoms to show that they are also subject to particle-wave duality and interference as well, and that the interference pattern can also disappear without disturbing the atom's path in any way. Furthermore, this will show that quantum erasure and delayed choices have the same effect on the system as predicted by QM for photons.

SEW built atom interferometers with detectors that were constructed with the aid of quantum optics devices, which emerged from new technological advances of that time. They published their experiment in the renowned journal Nature under the title "Quantum optical test of complementarity". [7]

Fig. 12 shows a plane wave of incoming atoms from the left. They are collimated to produce a couple of well-defined beams of caesium atoms. These can be excited to a higher energy level by a laser beam, as seen in the figure. The laser's photons carry energy that is absorbed in energy packets, which then allow the caesium atoms to acquire a higher energetic configuration.

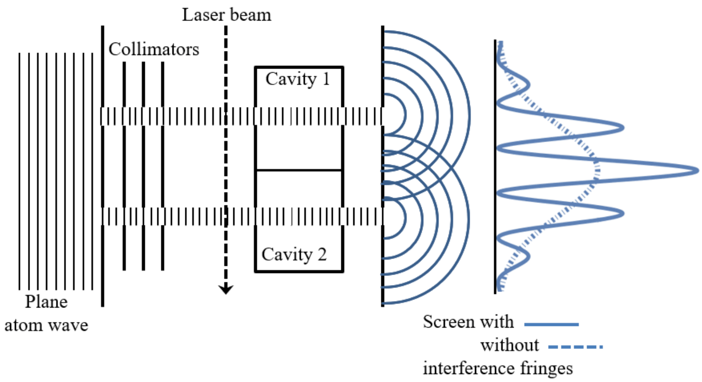


Fig. 12 Maser cavities for a which-way experiment with atoms.

The two streams of excited atoms are then sent into what are called '*micromaser cavities*'—that is, a couple of microwave cavities where EM radiation of a specific wavelength can be stored. The interesting point here is that the two micromaser cavities are capable of storing single photons. If the cavities are properly designed, they resonate for specific frequencies (that is, wavelengths or energies), they can store the photon emitted by a specific atom transition, and, if they are long enough, can even ensure that when an atom enters in an excited state it will emit its photon and leave the cavity in the ground state, with certainty. Therefore, the which-way information can be obtained through the act of reading out cavity 1 and cavity 2, to see which of the them contains the photon. Finally, immediately after leaving the two micromaser cavities, the atoms encounter the usual double slit and the detection screen.

Note that, contrary to what Feynman claimed, this is yet another example of how one can circumvent difficulties related to the uncertainty principle. While it is true that we impart a little kick to the atoms by illuminating them with the laser, this is really negligible if they absorb photons of low energy. And, anyway, the point is that the which-way information is not obtained by the scattering between the laser photons and the atoms but, rather, by controlling which cavity has stored the photon.

So, what will be observed when the laser is turned off? As in the case of the single photon counts in the double-slit experiment, here it is also possible to tune the atom flux in such a way that we send only one single atom at a time. When there is no laser beam, the atom is not excited, the micromaser cavities can't store the photon because the atom is in the ground state, and it can't release any photons when it traverses the cavities. Therefore, the cavities play the role of simply a further collimating device and we have no information to read out about the which-way the atom took. This implies that interference fringes will appear (the continuous line in the graph in Fig. 12) if we count a sufficient number of atoms, one after the other, hitting the screen. Effectively, this means that we are simply using a standard double slit device, but using atoms instead of photons.

What will happen when we turn on the laser? Then the interference fringes will be destroyed. This is because if an atom’s energy level is lifted up by a laser photon, it will release the photon with certainty in one of the two maser cavities, allowing us to determine which way it travelled along. Thus, the characteristic interference fringes will be replaced by the bell-shaped intensity curve (the discontinuous line in the graph of Fig. 12). This happens with almost no interaction with the atom beam, but only because of the contextual nature of QM. Recognize the analogy with the experiment of Fig. 9. There, the ‘slit-superposition’ of two orthogonally polarized single photon states led to the lack of interference fringes. Here it is the ‘cavity-superposition’ of the single atom symmetric and anti-symmetric wavefunction states (more on that next).

In a second experimental configuration (see Fig. 13), the two resonant cavities are no longer physically divided from each other, but are separated only by a pair of electro-optical shutters placed in front of a detector wall.

These can be closed or open—that is, let the photon through, or not, towards the common internal wall of ***both*** micromaser cavities. The latter is covered by a thin-film semiconductor which absorbs microwave photons and acts as a photodetector when the shutters are open. Let us consider what happens for the passage of a single atom in the two cases with the closed or the opened shutters.

We can send one atom through the system of cavities and the slits. Observe the individual spot appearing on the screen. Then, ***after*** the atom hits the screen, interrogate the micromaser cavities. That is, only after we see the spot on the screen will we make the delayed choice of whether we want to keep the shutters closed or open. This procedure is then repeated several times until an intelligible pattern can be recognized on the screen.

Note how we have built a delayed-choice device which furnishes the atom's which-way information but does not at all disturb the atoms themselves. We can let the atom hit the screen and then, only after that, control the state of the cavities that allow us to determine its path. Therefore, there could be no sort of physical interaction at all that could have destroyed the interference fringes. This is, again, another demonstration of Heisenberg's microscope fallacy.

This situation is analogous, though not entirely, to Feynman's microscope, which we discussed during the which-way experiments of Vol. I. There, an electron passing through the slits was entangled with a 'test-photon' shining on one slit in order to retrieve its whereabouts. Here, a caesium atom is entangled with the photon in both cavities.

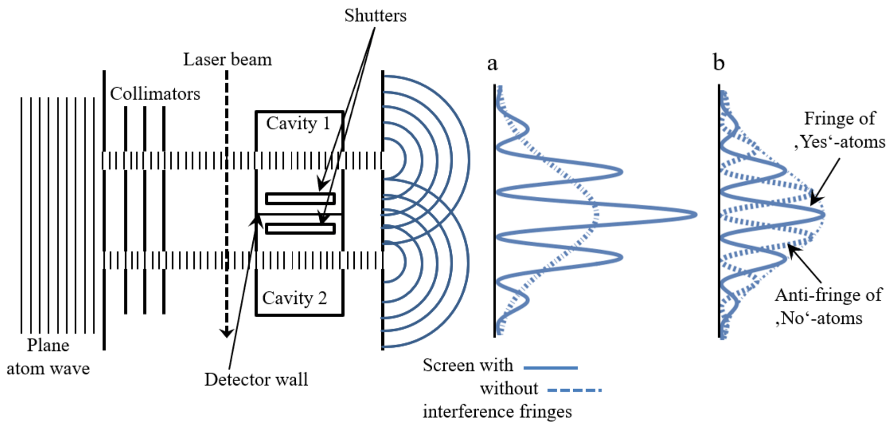


Fig. 13 Experimental setup of the SEW experiment.

Labeling the atom's path-state through cavity 1 and cavity 2 as $|\Psi_1\rangle_a$ and $|\Psi_2\rangle_a$ and labeling the photon being stored in cavity 1 and cavity 2 as $|1{,}0\rangle_\gamma$ and $|0{,}1\rangle_\gamma$, the entangled state between the atom and the cavities is:

$$|\Psi\rangle = \frac{|\Psi_1\rangle_a\,|1{,}0\rangle_\gamma + |\Psi_2\rangle_a\,|0{,}1\rangle_\gamma}{\sqrt{2}}.$$

If we chose to keep the shutters closed, the photon emitted by the atom in whatever cavity is not absrbed by the detector wall, and we can read out whether the atom left its photon in cavity 1 or cavity 2. Because the two cavities are placed in front of each slit, the apparatus with closed shutters is capable of telling us where an atom has gone through, by controlling in which cavity the photon has been stored. Therefore, we can extract information about the atom's which-way, which implies that no interference fringes can appear on the screen. (One again obtains the dashed line in graph (a) of Fig. 13).

If, instead, we chose to open the shutters and let the photon that the atom emitted during its passage in one of the cavities be absorbed by the detector walls, or simply be 'removed', then the 'memory of passage' (the which-way information) could be said to be 'erased'. That is, we have built a quantum eraser.

Consider that in this case (for quantum mechanical reasons too long to be discussed here), for an ideal photon detector having 100% efficiency, the probability that the detector wall will absorb the photon in both cavities is only 50%. (In the remaining cases, the photon remains unchanged and bounces back and forth in the cavity.) Let us label the atoms for the case of open shutters where the photodetector in the micromaser cavity clicked as a 'yes-atom', while those atoms where no photocount is observed in the cavities are called the 'no-atoms' and, respectively, the 'yes-' or 'no-eraser' photons.

If the choice of opening the shutters is made ***after*** the atoms hit the screen, the device acts like a delayed quantum erasure, because the atom leaves its photon in the cavity but without anyone reading out which of the two cavities, and then travels towards the screen and forms a spot. Only then do we decide whether, as in the previous case, we are going to keep the shutters closed and read out where the photon is, or open the shutters, whereby the photon is absorbed by the detector wall, erasing the information about its whereabouts. However, there is a 50% chance that the detector wall in the cavity will respond to the presence of the photon (the 'yes' eraser photon) and a 50% chance that it will not (the 'no' eraser photon). In both cases, we lose the which-way information because the no-photon is absorbed by the cavity walls after bouncing back and forth, and we will not be able to read out in which cavity the atom left it.

It turns out that if the which-way information is erased ***after*** the atoms hit the screen the interference fringes don't appear. This is because the atoms already hit the screen before the choice was made to erase the which-way information. To observe interference fringes, we must perform the quantum erasure before the atoms hit the screen. Otherwise, it is too late. It is like waiting to insert the $\pm$ 45° polarizers in Fig. 10 until after the photon hits the screen and still hope for the interference fringes to reappear. It is hard to believe that this could be the case, as that would imply a retro-causal action into the past. And, in fact, it isn't. Otherwise, it would appear that the particles or atoms detected in the present (future) time must inform their own 'selves' in the past (present) about whether they should take a path that forms a bell-shaped or fringy pattern. Another analogy that comes to mind is the MZI version of Wheeler's delayed choice experiment. (See the delayed choice experiment in Vol. I.) We showed that a delayed choice is possible while the photons are 'in flight' before the detection, between the first and second beam splitters in the MZI. If the quantum erasure process was performed during that short time period, you recover the interference fringes. However, if it was performed after it went through the second beam splitter, then it is too late; we will accordingly obtain the lump of particles on the screen.

However, this is only part of the story. As usual, Nature is tremendously subtle and is able to mix things up when the human mind can see only mutually exclusive, logical options. It turns out that one can recover the interference fringes from the collected data. After a suitable amount of time, during which you have collected a sufficient number of events (that is, several atoms hitting the screen), the spots on the screen will build up the classical Gaussian bell-shaped probability function. However, when you look only at the spots left behind by the 'yes-atoms', those where the detector clicked, or, alternatively, the spots correlating with the 'no-atoms', when the detector did not click (one can separate the two sets of spots on the screen, as one knows when the photodetector clicked for the yes-atoms and did not click for the no-atoms), then the good-old interference fringes are recovered. The continuous and the dashed interference fringes in the graph (b) of Fig. 13 show the two cases. The fringes are equal but shifted interference patterns comprising the fringes for the 'yes-atom' and the anti-fringes for the 'no-atoms'. When we sum them up, the bell-shaped curve appears and the interference pattern disappears—or, more precisely, remains hidden.

So, is there a contradiction in the temporal order of events? Fringes and anti-fringes are there, aren't they? The delayed quantum erasure process was performed after the atoms hit the screen. We can somehow recover the interference fringes by correlating the 'yes-' and 'no-atoms' with the spots on the screen. This seems to suggest that the atoms chose to displace themselves on the screen according to a choice that still had to be made at the time they hit the screen! How can an interference pattern appear if the physical process that is supposed to determine it (our choice of opening the shutters erasing the which-way information) is performed after it came into existence?! How could the 'yes-atoms' and 'no-atoms' know that we would have erased the which-way information and distribute themselves on the screen according to an interference pattern when this choice had yet to be made at that time? We are apparently again confronted with quantum retro-causal effects: The choices we make in the present determine the atoms' past behavior. Or, if you prefer, present physical events are influenced by the future. The behavior of a quantum system today seems to depend on events that will materialize tomorrow.

The author would love for that to be true, as this would, in fact, tend to confirm a symmetry in which past, present, and future are one. This is a vision of things with which he is sympathetic. However, a healthy skepticism combined with factual objectivity—and not personal or ideological preferences— should lead our thoughts.

Everything becomes intelligible when we realize that the 'yes-' or 'no-atom' correlation is decided at the time when the atom hits the screen, ***not*** when we open the shutters. What we do is no 'delayed choice' at all; rather, it is just a control, a readout of the whereabouts of a photon that had already come into existence before the choice. The point is that the state vector represented by a superposition of states (the atom taking both ways with the photon being stored in both cavities at once) undergoes a state reduction, that is 'collapses', only at the instant when the atom hits the screen. It is only at this point that the state vector is projected onto one of the two possible eigenstates (the atom taking only one or the other path and the photon stored in one or the other cavity, but not both). Then the 'game is over'.

In fact, when the atom hits the screen, it sets the wavefunction of the photon inside the cavity to a symmetric or anti-symmetric state by a non-local correlation. Then, this radiation remains inside one or the other cavity (but not both) and couples to the photon-counter detector wall of the micromaser cavity only later, when one opens the shutters (but not before; we do that after the atom hits the screen). It is the former or the latter wavefunction, the symmetric or anti-symmetric wavefunction, which determines whether or not the detector will fire. Therefore, before our decision regarding whether to open the shutters, the photon's destiny is determined at the instant of the atom's detection on the screen, not by our choice. So, there is no problem with retro-causations here.

As a side note, you might have noticed a similarity between the two interference fringes of the Aharonov-Bohm effect (see Vol. I), those in which the magnetic field in the solenoid behind the slit screen is turned on or off, the neutron interferometer experiment and the interference fringes of the yes- and no-atoms in this SEW experiment. In the former case, the shift of the peaks of the fringes was due to a phase difference that the magnetic potential vector imparted on the wavefunctions corresponding to the two states with the magnetic field on or off. Here, the same happens but with the phase difference due to the 'yes' or 'no' states.

The bottom line of this experiment is the following. Composite material systems like atoms (or, eventually, even molecules, as shown in other experiments not discussed here), and not merely photons or elementary particles, are also subject to the wave-particle duality. The SEW experiment shows that these quantum phenomena are not due to the interaction or perturbation with the atoms; it is an intrinsic property, a law of nature, that the wave or particle character realizes in accordance with the specifics of the situation. This also leads to the conclusion that complementarity must be a universal and very fundamental feature for all particles and systems of the quantum world. Heisenberg's uncertainty principle is only one aspect of it, and is not the cause or source of complementarity. Therefore, complementarity is more fundamental than the uncertainty principle. Finally, no mystical retro-causation is needed to explain the facts. We must always keep in mind the difference between the evolution of the wavefunction and its collapse. Once the latter has occurred, the state reduction reflects itself non-locally throughout the system.

4. The delayed quantum erasure experiment of Walborn et al.

About then years later, in 2001, yet another overall combination of a double-slit delayed quantum eraser experiment was performed by a Brazilian group (Walborn et al. [8]). It might be instructive to dwell further on these experiments to overcome doubts about any supposed retro-causality in QP. It can also be considered the photonic replica and continuation of the SEW experiment.

Fig. 14 illustrates how photons from an argon laser (at 351 nm wavelength) are focused by a lens and sent through a BBO crystal to create, via a SPDC, a couple of entangled photons.

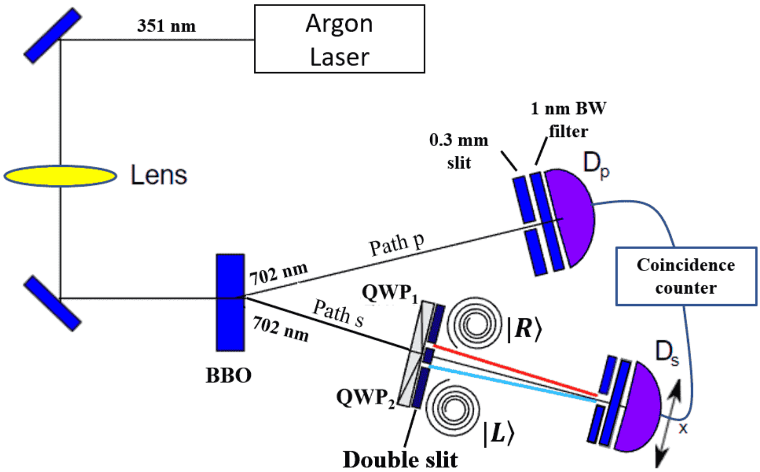


Fig. 14 The which-way experiment with entangled photons: version I.

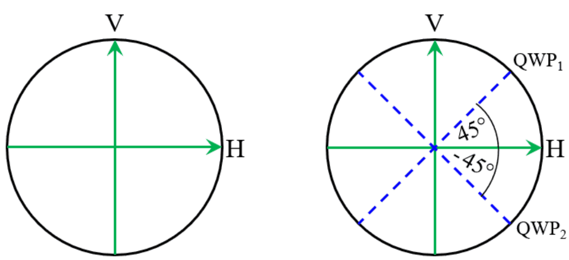


Fig. 15 Polarization diagrams for experiment version I.
Left: for path p. Right: for path s.

Because the total energy must be conserved, the wavelength of the two down-converted photons must be twice as much as the original one (or equivalently, they have half the frequency of the incoming photon, of course, due to Planck's relation between the energy and frequency). Therefore, the two photons have a wavelength of 702 nm (deep red color). They are orthogonally polarized—that is, they are type-II entangled photons (in contrast to the type-I having the same polarization, see the discussion in Vol. I on SPDC entangled photons). This means that, as long as we do not measure it, we must conceive of it as being in both horizontal (H) and vertical (V) polarizations at the same time, as shown by the solid arrows in the polarization diagram of Fig. 15. (More generally, an orthogonal polarization considers any 90° basis, not only horizontal-vertical orientations. However, to keep things simple, we restrict ourselves to the 45° and -45° basis.) The photon travelling the upper path (let us label it path 'p') is sent towards a small slit of 0.3 mm, behind which a 1 nm band width (BW) filter is placed (to ensure that no other photons from the environment will be counted). It is then detected by a photodetector D_p.

In a brief interlude, without going too much into the technical details, let us also mention how circular polarized light can be obtained from unpolarized light. This can be done through us of a '*quarter-wave plate*' (QWP), as shown in Fig. 16. From the right to the left: Unpolarized light is first rendered linearly polarized with a linear polarizer and then filtered again with a QWP, which converts it into circularly polarized light. A QWP is a retardation sheet such that horizontally and vertically polarized light entering in phase will emerge from the retardation plate at 1/4 of a wavelength ($\frac{1}{4}\lambda \sim 90°$), out of phase. If linearly polarized light enters at an angle of $+45°$ between the fast and slow axis, then the x and y components of the electric field will be phase shifted with one component lagging behind, resulting in a rotating electric field vector, that is, a left-handed circular polarization. Whereas, a $-45°$ linearly polarized light (not shown in the figure) results in right-handed circularly polarized light.

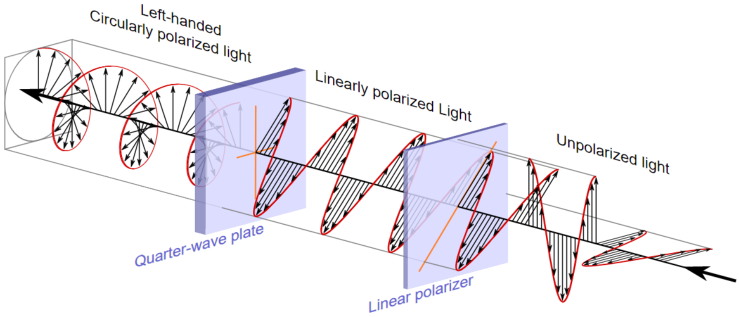


Fig. 16 Circular polarization of light.

From the quantum mechanical perspective, this amounts to saying that the photons emerging from this polarization device possess a specific spin, that is, a positive or negative helicity. (See also the chapter on the photon's polarization and spin in Vol. I.) QWPs are not polarizers which select photons with a specific polarization; rather, they are devices that change the polarization of all the incoming photons, ideally without absorption.

Now apply QWPs in this context: The entangled photon following the lower path, labeled ‘s’, which is also in the superposition of the horizontal and vertical polarizations, before going through a double-slit device, will encounter two circular quarter-wave plates, QWP_1 and QWP_2, placed in front of each slit— say, QWP_1 in front of slit S_1 and QWP_2 in front of slit S_2. (By analogy, compare this to the experiment of Fig. 9.) What differentiates the two QWPs is their orientation (relative to H and V of their fast and slow axes), which is shown by the dashed lines in the polarization diagram of Fig. 15 right. A +45° or -45° difference between the photons’ polarization and the QWPs orientation always leads to an opposite circular polarization. Therefore, QWP_1 in front of slit 1, will impart to the photon a left-handed (clockwise) circular polarization $|L\rangle$ or, equivalently, a negative helicity — that is, the photon has $S_\gamma = -\hbar$ spin. Meanwhile, QWP_2 in front of slit 2 will impart it a right-handed (counter-clockwise) circular polarization $|R\rangle$ or, equivalently, a positive helicity—that is, the photon has $S_\gamma = +\hbar$ spin. Photons will then also traverse a small slit behind which another 1 nm filter is placed and will then be detected by photodetector D_s. A stepping motor moves detector D_s, scanning along the x-direction in order to read out the intensity curve with a photocounter. Overall, only those photons are counted when a coincidence counter confirms the detection of both photons s and p.

With this experimental setup, we have built a which-way tester. In fact, the role of the two QWPs inducing circular polarization in front of the two slits should not have gone unnoticed: The QWPs ‘mark’ the photon. No matter what happens to the photon on path p, this allows us to determine the which-way. It shouldn’t come as a surprise that no interference fringes will be measured at D_s and that one observes only the usual bell-shaped intensity curve as shown in Fig. 17. (The vertical bars represent the measurement error).

From the wave perspective one must consider how the clockwise and counter-clockwise polarization vectors rotate. Their angular difference $\delta\theta$ varies very rapidly (with the frequency of a 702 nm light wave!) between 0° and 180°. That is, the interference patterns of Fig. 7 change and mix together extremely fast so that the net resulting interference pattern is the sum of all these, leading to the Gaussian probability function.

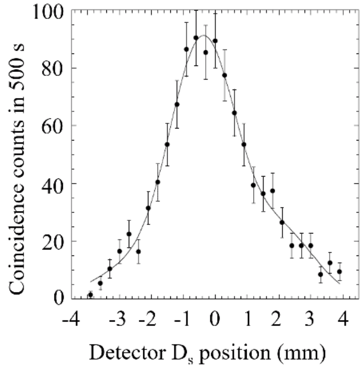


Fig. 17 Coincidence counts at D_s with QWPs on path s but without $POL_{1/2}$ on path p.[8]

In a second version of this which-way experiment (see Fig. 18), one inserts in the upper path, that of photon p, a linear polarizer $POL_{1/2}$ in front of detector D_p. If it is oriented along the 45° direction we label it POL_1, while for the −45° tilt it is labeled POL_2. Keep in mind that, again, a polarizer selects photons only. There is a 50% probability that POL_1 (POL_2) will allow the photon in a vertical and horizontal polarization superposition to slip through as a 45° (−45°) photon, while the others are blocked. That is, only half of the photons on path p will be detected by detector D_p, it will see only those photons. Meanwhile, the entangled photon on paths s will always hit detector D_s, which will observe, instead, a stream of photons with both possible polarizations. However, this is of no concern because what we are looking for are the coincidental counts where both photons are detected and correlated to each other. We are not looking for the interference pattern they produce individually at each detector.

Another important aspect to notice is that $POL_{1/2}$ is placed, from the BBO, at a distance about half as long as the distance of detector D_s, but farther than the slits. This implies that the first detection will always take place at D_p when the photon on path s has already traversed the slits but is still on its way to detector D_s.

Let us analyze this second case in more detail. Recall how the measurement of entangled photons works. For type-II entangled photons, if on path p a linear polarizer selects a photon along the H (V) polarization direction, the photon on path s must anti-correlate and be in polarization state V (H). When the evolving state function is before the two slits, it must be represented by the entangled state, which we know to be:

$$|\Psi\rangle = \frac{|H\rangle_s|V\rangle_p + |V\rangle_s|H\rangle_p}{\sqrt{2}}, \quad \textit{Eq. 6}$$

with the obvious labeling 'p' and 's' being the photon on path p and s respectively.

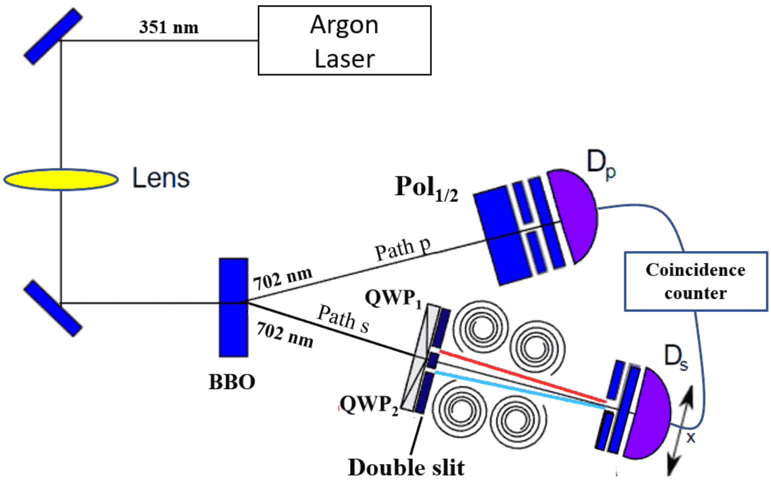


Fig. 18 The which-way experiment with entangled photons: version II.

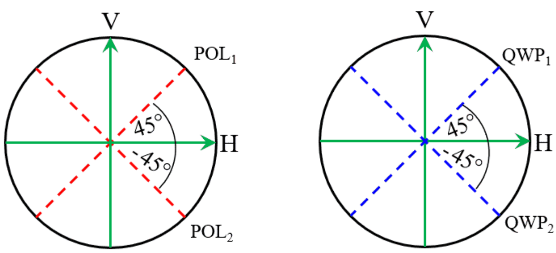


Fig. 19 Polarization diagrams for experiment version II.
Left: for path p. Right: for path s.

Once the photon on path s has traversed the two slits, it is set into 'slit-superposition' S_1 and S_2 and its linear polarization is transformed into the circular polarization. Remember that this is not a measurement and therefore there is no state collapse. It is the state function that has been changed by the slits and circular polarizers without any interaction or information readout.

Each photon on path s emerging from slits S_1 ***and*** S_2, having either vertical or horizontal polarization, is also entangled with photon p. (One might say that there are three beams but only two particles!). There is polarization and slit superposition intertwined with entanglement. Therefore, the overall system can be described with the state vector $|\Psi\rangle$ as:

$$|\Psi\rangle = \frac{|\Psi\rangle_{S_1} + |\Psi\rangle_{S_2}}{\sqrt{2}}, \quad \textit{Eq. 7}$$

where

$$|\Psi\rangle_{s_1} = \frac{|H\rangle_{s_1}|V\rangle_p + |V\rangle_{s_1}|H\rangle_p}{\sqrt{2}} \quad \text{and} \quad |\Psi\rangle_{s_2} = \frac{|H\rangle_{s_2}|V\rangle_p + |V\rangle_{s_2}|H\rangle_p}{\sqrt{2}}.$$

Expressing it in words, it means that the state of the system is in a superposition of four potentialities: Photon s goes through slit 1 with horizontal polarization, with photon p having vertical polarization AND also with a vertical polarization with photon p having horizontal polarization AND the same photon s also goes through slit 2 with horizontal polarization, with photon p having vertical polarization AND also with vertical polarization, with photon p having horizontal polarization. Insertion of the two QWPs in front of the two slits, as in the second experimental version, would replace the horizontal and vertical polarizations with the left- and right-hand circular polarizations (with the addition of some subtleties that we do not discuss further here). Maybe this clarifies why mathematical formalism is so much more useful than employing annoying, lengthy, and clumsy sentences!

At this point, we can proceed on the same line of discussion that led us to Eq. *5*. There, a photon before entering the slits is in a diagonal polarization state and is then filtered by a horizontal and vertical polarizer after the two slits (see Fig. 10). The outgoing evolving quantum state could be expressed in a diagonal polarization base as a superposition of a two-slits-state with a symmetric and anti-symmetric wavefunction. In this experiment things are pushed further than that, as to the two-slits-superposition of the photon on path s one must also add the entanglement with the photon on path p. Here also, the orthogonal states $|H\rangle$ and $|V\rangle$ on paths s or p and the rotating polarization vectors $|R\rangle$ and $|L\rangle$ after slits S_1 or S_2 can be expressed in a $\pm 45°$ basis as a superposition of a symmetric and anti-symmetric wavefunction. The principle is similar to that which furnished Eq. *5* (though algebraically more involved and we won't develop it here, the interested reader is referred to the original article [25]) and leads to the following sate function:

$$|\Psi\rangle = \frac{|45°\rangle_{S_1} - i|45°\rangle_{S_2}}{2}|45°\rangle_p + i\,\frac{|-45°\rangle_{S_1} + i|-45°\rangle_{S_2}}{2}|-45°\rangle_p. \quad \textit{Eq. 8}$$

The analogy with Eq. *5* is manifest except for the imaginary numbers which account for a phase difference of 90° induced by the QWPs fast and slow axis on the two slit paths and by exchange of the fringes with anti-fringes. It contains the two slits-superposition states on S_1 and S_2 of the photon on path s times the entanglement with its twin photon on path p. The two terms express an anti-symmetric and symmetric state, respectively. On the line of Eq. 7, these represent not just the state of one or the other photon, or the superposition of one photon, or the entanglement of two photons, but all that put together as an overall quantum state describing a unique and inseparable whole as an entanglement between paths p and s plus a slit/polarization-superposition on path s. Any inference about ‘two individualized photons on two paths, one of which is going through one or the other slit’ exists only in our fantasy and has nothing to do with reality—at least, as long as state projection does not occur.

Note that the very same state function is equivalent to place polarizer $POL_{1/2}$ after the QWPs instead on path p. In doing so, one would insert a quantum eraser, as a $\pm 45°$ filter would pick out all the photons along that angular direction and render the two circular polarizations indistinguishable—that is, it erases the marking circular polarization which previously allowed for which-way information and the reappearance of the interference fringes. Filtering out the wavefunction after the slits and the QWPs with a 45° polarizer selects the left-hand side anti-symmetric state function of Eq. 8 and displays an interference anti-fringe pattern. Doing so with a -45° polarizer would select the right-hand side symmetric part of Eq. 8, and the fringes would appear along the x-direction of detector D_s.

However, no diagonal polarizer is inserted along path s but, rather, on path p! It is by orienting the polarizer along path p, as POL_1 or POL_2, that one can select the symmetric or anti-symmetric part of the wavefunction emerging from the slits. We might say that entanglement allows not only for ‘spooky actions at a distance’ but also for a sort of non-local instantaneous ‘insertion at a distance’ of a polarizer along a path without its physical presence there.

In fact, if the polarizer on path p is set into a 45° orientation and the photon in the $|H\rangle_s$ and $|V\rangle_s$ superposition will get through (50% chance), then, once it reaches detector D_p, state collapse occurs and by a ‘click’ we know that it must have been a $|45°\rangle_p$ state photon. And because, before the collapse occurring at D_p, it was entangled with the photon on path s, which has already traversed the silts, the photon will be set on path s into the anti-symmetric quantum state and the 45° polarization state. This implies that the quantum erasure is performed when the photon is in flight between the slits and D_s. Once the photon on path s hits detector D_s it will displace itself on one of the anti-fringes. Because only one photon at a time travels through the entire experimental setup, if the photon on path p made it through POL_1 with a 50% chance, then the coincidence counter will register how both detectors clicked (even though not at the same time) and correlate the photon on path p with polarization $|45°\rangle_p$ with the spot on the anti-fringe pattern along the scan of detector D_s. Similarly, if the polarizer on path p is set into a -45° orientation and the entangled photon will get through (a 50% chance as well), detector D_p will also click but we know that it must have been a $|-45°\rangle_p$ state photon. And, again, once it collapses at D_p, it will set photon on path s into the symmetric -45° polarization quantum state which will displace itself on one of the fringes of detector D_s. The coincidence counter will correlate the photon on path p with polarization $|-45°\rangle_p$ with a spot on one of the fringes. These two cases are summarized in the coincidence count of Fig. 20.

So, this is a three-step process: First, the $POL_{1/2}$ selects one of the two sides of Eq. 8 but does not collapse or 'lift' either the entanglement between the two photons or the 'slit-superposition state' of the photon on path s. However, at this stage, it has already been decided whether the photon on path s is in a symmetric or anti-symmetric quantum state. Secondly, the first state reduction will occur only once the photon on path p reaches detector D_p. Once this has occurred, the photon on path s is no longer entangled but is still evolving in a (symmetric or anti-symmetric) slit- superposition state. It was the polarization and subsequent measurement of the photon on path p that 'steered' the quantum state of the photon on path s. One speaks of *'quantum steering'* (concept introduced by Schrödinger) when for two entangled systems, the quantum state of one system can be prepared by a measurement on the other. Third, the photon on path s also reaches detector D_s. The second state reduction occurs, and the game is over.

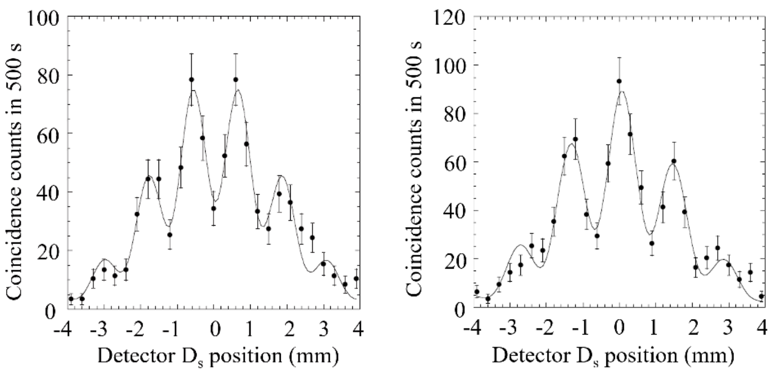


Fig. 20 Coincidence counts at D_s with POL_1 or POL_2 on path p.[8]

Then, the overlying of the two fringe and anti-fringe interference patterns will result in the Gaussian bell-shaped curve. This is in analogous to what we have elucidated at length with respect to the experiment of Fig. 10, in which the two orientations of the polarizer led to a fringe and anti-fringe pattern respectively and their sum to Fig. 11. In fact, if we would look directly at the pattern emerging on D_s without separating the fringe from anti-fringe photons by means of the coincidence counter correlation, we would observe only the Gaussian bell-shaped curve without a sign of any interference. It is only by noting the 'coincidence click'—that is, by correlating the $|45°\rangle_p$ or $|-45°\rangle_p$ photons on path p with the fringe or anti-fringe photons on path s respectively—that we can filter out the fringe from the anti-fringe pattern in a figure that otherwise would result in a normal distribution. And how could it be otherwise? Locally, there is no polarizer $POL_{1/2}$ in place on path s and from the point of view of the detector D_s, the photons going through the two slits and the two marking circular polarizers are still distinguishable. It is only with the entanglement between the two photons and the quantum steering action at a distance of $POL_{1/2}$ that one can act 'as if' the same polarizer is placed on path s and transform the evolving state function such that the photons traveling towards D_s will 'lose their marking' and behave like waves. But, at D_s, we have no knowledge of whether the entangled photon on path p collapsed into a $|45°\rangle_p$ or $|-45°\rangle_p$ state. It is only the correlation (which must be communicated via a classical communication channel) that reveals which was in which state.

So, the bottom line is that the quantum erasure induced by a distant polarizer via entanglement with a non-local action on a which-way experiment indeed allows for the reappearance of the interference fringes, as well as when photons on path s have already passed the silts and the marking of circular polarizers. But this 'recovery' comes at a price: These fringes and anti-fringes are subsumed and hidden in the bell-shaped curve and can be filtered out only by a count that correlates the photons' state on the two paths. This is perfectly analogous with the SEW experiment in which the (anti-)symmetric (no-)yes-atom wavefunction led to the (anti-)fringe patterns.

The natural question that arises at this point is: What would happen if polarizer $POL_{1/2}$ is placed farther along path p (eventually even light years away)? Say that the distance of D_p from the BBO crystal is much greater than that of D_s, as illustrated in Fig. 21.

This implies that the quantum erasure is delayed. One speaks of a '*delayed quantum erasure*' (not to be confused with a which-way delayed choice!) because photon p is set into a linear polarization state and absorbed by detector D_p only after (eventually years after) photon s has already reached detector D_s and has been absorbed by the measurement process. Moreover, here, the 'choice' is not made by a human observer manipulating polarizer $POL_{1/2}$. Rather, it is decided by Nature's random determination of which of the two possible states arises after the entanglement is gone due to the measurement at D_s first.

The answer that might sound initially surprising is that the order of detection—that is, whether detector D_s clicks first or detector D_p clicks first— is not relevant. As long we correlate the photons on the two paths by the coincidence counter, the interference fringes and anti-fringes remain there even if polarizer $POL_{1/2}$ is far away and the quantum erasure is delayed until photon s hits detector D_s. This might suggest, again, that a sort of retro-causality exists whereby the $|45°\rangle_p$ or $|-45°\rangle_p$ state of photon p in the future determines the symmetric or anti-symmetric state of photon s in the present. Or, equivalently, that 'lifting' the entanglement state of photon p in the present retro-causes the fringe or anti-fringe displacement of photon s in the past. This sounds extremely weird and is reminiscent of the time machines of sci-fi fantasy movies.

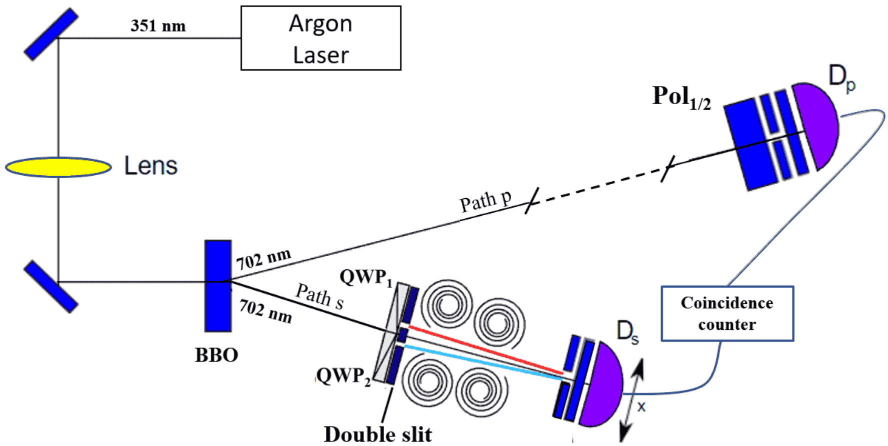


Fig. 21 The which-way experiment with entangled photons: version III. The delayed erasure: as in version II but with D_p shifted along path p.

Again, the analogy with the SEW experiment is straightforward. The delayed quantum erasure in this third experiment version, in which detector D_p is much farther from the BBO than detector D_s, corresponds to the situation of the SEW experiment with the opening of the shutters at a later time than the time of absorption of the caesium atom on the screen. A clumpy curve of particles appeared.

Things are similar here, too, and much more simple and down to earth than any 'back to the past actions'. There is no need to invoke any retro-causality if we remember where and how the collapse of the state function occurs and especially where it does not. In the first case, with detector D_p clicking first, the state collapse occurs only for the entanglement of the system, with the two photons acquiring a separate existence. However, the two-slit superposition of the photon on path s remains unaffected. On the other hand, once the photon on path s hits detector D_s a state reduction is caused that involves the superposition as well. At the instant of detection on detector D_s, the photon on path p will be set into a corresponding diagonal definite quantum state $|45°\rangle_p$ or $|-45°\rangle_p$ already before it hits the (light years away) detector D_p. As in the previous experimental version, polarizer $POL_{1/2}$ doesn't collapse anything, it simply filters out the photon in a $|\pm45°\rangle_p$ state.

Remember that only those photon correlations between those on path p and paths s are counted when the coincidence counter registers the correlation. If a photon on path p is blocked by polarizer $POL_{1/2}$ (as in 50% of the cases), detector D_p does not measure its presence and these entangled photon pairs are discarded from the data count. Finally, we 'uncover' only a correlation. The click of detector D_p tells us only which photon to pick out and ascribes it to the fringe or anti-fringe photons. However, it does not retro-cause anything. This is one of the few instances in which QP is less weird than it might appear. There is nothing 'mystical' about the delayed quantum eraser and the classical temporal order is safe. To be more precise, in this second configuration of the experiment, there is no 'delayed erasure' at all. The 'game was already over' when the photon on path s was absorbed by detector D_s. It would be misleading to suggest that quantum retro-causality exists, as some have done.

In a certain sense, the first part of the experiment is much more interesting, when the selection of polarizer $POL_{1/2}$ and the collapse at detector D_p takes place before the photon on path s hits detector D_s. In that configuration, QM reminds us, again, that as long as the state vector projection has not taken place, we must not regard a quantum system as being made of different separate and independent parts. Rather, we must regard it as a unique and undifferentiated non-locally connected whole. Eq. 7 must be taken seriously, not just as an abstraction without reality. It is only by adopting this perspective that we can understand how, once a measurement is taken of one element of the system, this reverberates on the whole system via an instantaneous non-local 'action'. The state of a quantum system must be conceived of as a potentiality that realizes itself with some probability and that describes the system as a whole, never as a local subset of parts put together but, rather, as a non-local oneness that only upon the act of measurement (state collapse, reduction, projection) distinguishes among the several possible potentialities. If we stick with the idea that a separation exists between beams and particles going through slits or polarizers collapsing wavefunctions, we will be dangerously prone to the separation and measurement fallacy.

5. Putting it all together: the delayed choice quantum eraser of Kim et al.

Almost simultaneous with the experiment of Walborn et al., another fascinating DCQE experiment was performed by a group of scientists from the University of Baltimore and the Texas A&M University (Kim et al. [9]). This experiment did a good job of putting everything together: the double slits, quantum entanglement, a delayed choice quantum erasure (well, not really as we will see), and an interaction free which-way measurement. A sketch of the experimental setup is shown in Fig. 22.

An argon pump laser beam (351 nm wavelength, ultraviolet light) shines on a double slit behind which a BBO crystal generates a pair of entangled type-II orthogonally polarized photons (702 nm wavelength each) by SPDC at regions A and B (0.7 mm center separation). As usual, the light beam is a single-photon source. The signal-photon, that which goes through the lens focusing on detector D_0, and the idler-photon that which travels towards the prism. Detector D_0 scans with a step motor along the perpendicular axis of the path of the incoming signal photon (as shown by the arrows at D_0 in Fig. 22) counting each photon cumulatively in order to reconstruct the interference pattern (that is, the fringes or the bell-shaped intensity curves).

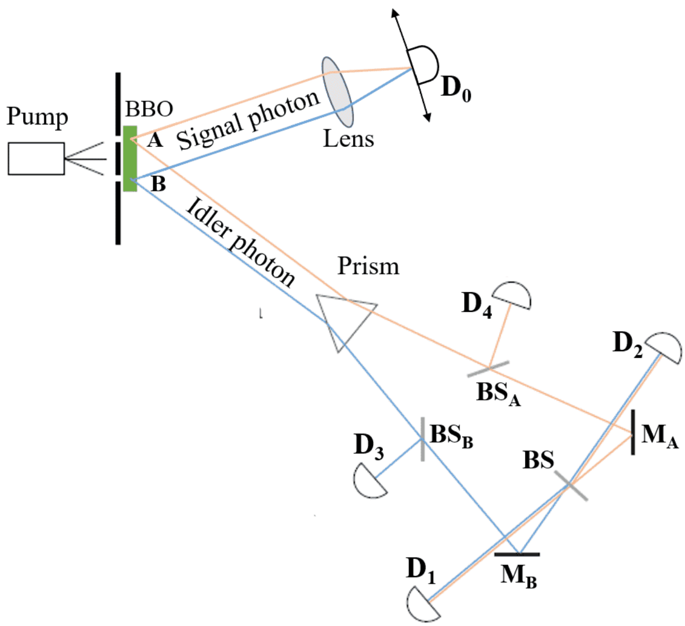


Fig. 22 The delayed choice quantum eraser of Kim et al. [9]

The prism (a *'Glan-Thompson prism'*) has the peculiarity of splitting orthogonally polarized beams. Because regions A and B determine two mutually orthogonal polarizations, the splitting prism sends the idler-photon towards beamsplitter BS_A if it has a horizontal (vertical) polarization or towards beamsplitter BS_B if it has a vertical (horizontal) polarization. Therefore, in the event (50% chance) that this photon, according to its polarization, is deflected towards detectors D_3 or D_4, one can determine the which-path from region A or B, respectively. However, if, instead, it is transmitted through beamsplitters BS_A or BS_B, it will travel farther and be reflected at mirrors M_A and M_B and, later on, will encounter a third beamsplitter (BS). Here, again, it can be either reflected or transmitted with a 50% chance. If it is coming from (or, more precisely, we imagine that it is coming from) mirror M_A (M_b), a reflection at beamsplitter BS sends it to detector D_2 (D_1), whereas, in the case of transmission, it will travel towards D_1 (D_2). This means that whenever detector D1 or detector D2 clicks, we can no longer determine whether it came from region A or B, because for both cases there is a 50% chance that one or the other paths has been taken. In other words, beamsplitter BS works as a (passive) quantum eraser of the which-way information.

Finally, a coincidence circuit (not shown in Fig. 22.) correlates each photon measured along the stepping-axis of detector D_0 by a 'joint detection' on one of the four detectors that its twin photon has triggered.

An important aspect of this experimental configuration is that it has been built in such a way that the optical path of each idler-photon—whichever path it will take from the BBO to whichever detector $D_{1\text{-}4}$— is at least 2.5 m longer than the optical path of the signal-photon from the BBO to detector D_0. This means that detector D_0 is always triggered first and that only later (by a delay of at least 8 ns) will one of the four other detectors click. This space-like separation is a necessary condition for ensuring that there is no potentially unknown physical effect that might 'inform' detector D_0 about what the other detectors will do in the future. If it does, then only FTL effects could be responsible, but no causal correlation that the theory of relativity allows for. Of course, an extremely fast electronic readout at the detectors is necessary and must occur in a time-lapse no longer than a few billionths of a second. This is, however, no issue for modern electro-optical devices.

Once you have this experimental configuration clearly in mind, it should not be difficult to see that:

a) If detector D_3 (D_4) clicks, the idler- and signal-photons could have come from only slit B (slit A)— that is, we have delayed which-way information.

b) If detector D1 or detector D2 clicks, the idler- and signal-photons could have come from either A or B—that is, we have a delayed which-way information erasure at beamsplitter BS.

The measurements are performed by collecting the data of the 'joint detection rate', R_{0j}, between detector D_0 and $D_{1\text{-}4}$. The detection time at D_0 and $D_{1\text{-}4}$ is not exactly simultaneous because of the mentioned minimum 8 ns delay that temporally separates the incidence of the signal-photon at D_0 and the idler-photon on one of the other four detectors. However, if the coincidence monitor registers a common detection at D_0 and $D_{1\text{-}4}$ within a time interval no longer than that which light needs to traverse the entire experimental setup (the single-photon source provides that no other photons are 'in-flight' during that time lapse), this confirms the coincidence as a joint detection, correlating the individual signal-photon at D_0 with the idler-photon at $D_{1\text{-}4}$. Moreover, the coincidence monitor allows for the filtering out of unwanted random photons coming from the environment, which otherwise would corrupt the data set as noise.

So, what should we expect to see? First, let us take a look at detector D_0 in front of the double slit and scan its optical focusing plane to reconstruct the overall intensity profile considering ***all*** the incoming signal photons.

The answer is straightforward: Despite the fact that D_0 cumulatively measures the photons emerging from two interfering waves from slits A and B, no interference fringes will appear; only the clumpy Gaussian curve is observed. Due to the generation of orthogonal polarizations in front of slit A relative to slit B, this 'marks' the photon and we have the which-path information from which slit we imagine it to have come.

However, with the background knowledge we gained from section I.2, we can also reach the same conclusion using classical wave optics, without the need to invoke the 'quantum which-way fiction'. We know that two orthogonally interfering waves leave no trace of any double slit interference fringes, as was amply explained in chapter (see Eq. 1, Fig. 7 for $\delta\theta$ =90° or Fig. 9 or Appendix A II).

The next question is: Which idler photon triggering the other four detectors correlates with a 'joint detection' of the signal photon detected at D_0?

Let us first consider the joint correlation rates for D_3 and D_4. These detectors measure the idler-photon (filtered by the prism and reflected by beamsplitters BS_A or BS_B), which is vertically or horizontally polarized—that is, if it is coming from slit A or B, and therefore 'revealing' the which-way. It is therefore no surprise that, if one were to place a detection screen at D_3 or D_4, one would again observe no interference fringes. After all, how can there be any? The two beams have been separated from each other and we are looking at one or the other slit separately, preventing the two waves from interfering in the first place. (Recall a similar situation in Wheeler's delayed choice experiment in Vol. I with no screen in place and only the two detectors pointing at the slits.) Which polarization is assigned at which photon from which slit is a completely quantum random process that is determined at the instant of the collapse of the idler photon. That is, at each new photon coming from the source, the horizontal and vertical polarization is assigned randomly between path A and path B. The idler photon's polarization will continuously switch sides. Despite the anti-correlation with the signal photons (always keep in mind that the BBO produces type-II entangled photons), this does not determine whether they displace themselves on a fringe or anti-fringe of the interference pattern. Therefore, detectors D_3 and D_4 will reproduce the same intensity curve that the signal photons produce on detector D_0 (with one half of the intensity, because they 'filter out' each 50% of the idler photons on average) and the coincidence joint detection rate, R_{03} and R_{04}, will be that of a bell-shaped curve as well (as shown in Fig. 23).

What about the joint detection rates between detector D_0 and detectors D_1 and D_2? First, note that if one were to place detection screens in front of these detectors, we would again observe the absence of interference fringes. This is because what both detectors D_1 and D_2 'see' is a superposition of a vertically and horizontally polarized idler photon. In fact, follow the paths of the idler photon in Fig. 22. One is coming from path A (B) and is reflected twice at mirror M_A (M_B) and at beamsplitter BS, while the other along path B (A) is reflected at mirror M_B (M_A) and transmitted through the same beamsplitter BS, both reaching detector D_2 (D_1).

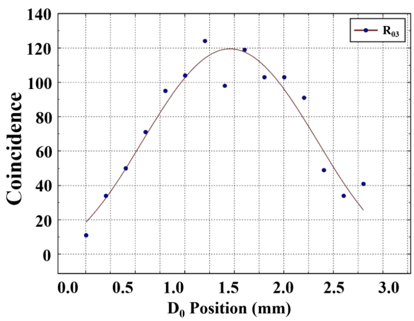


Fig. 23 Joint detection rate R_{03} (Kim et al. [9]).

The absence of interference here is somewhat surprising. This is because, if the second part of the experimental device comprising the two mirrors M_A, M_b, beamsplitter BS, and detectors D_1 and D_2 are supposed to work as a delayed which-way quantum eraser system, we should expect to see interference fringes at detectors D_1 and D_2. Indeed, a closer inspection reveals that this is not really a quantum eraser, at least not in the conventional sense. Suppose the idler is in a horizontal polarization state on path A and a vertical one on path B. Then, if we were to place a polarizer in front of detectors D_1 and D_2 (say, both with a horizontal polarizer), then if D_1 or D_2 clicks, the idler photon must have traveled along path A. If none click, the idler photon must have been a vertically polarized idler photon traveling along path B. So, in a certain sense, this is a 'fake which-path quantum eraser'. What differentiates detectors D_1 and D_2 from detectors D_3 and D_4 is that the latter 'see' the single slit by pointing at it directly and already determine, in advance, which path they are measuring, independently from 'polarization marks', while the former 'see' both slits at once. However, for anti-correlated photons, this does not prevent us, at least in principle, from determining the path using analyzing polarizers.

The catch is that, as already mentioned, we never know which idler photon has which polarization on which path. The BBO produces anti-correlated entangled photons, but which polarization is associated with which slit remains a quantum random event. The indistinguishability arises due to the random polarization labeling of each idler photon or, if you prefer, the random switching between slits A and B. It is this continuous quantum random side switching that prevents us from knowing the which-path. It is not the beamsplitter that makes the paths indistinguishable.

The question, then, is: Is this a quantum eraser or not? If we regard the existence or the lack of interference fringes as the ultimate test for the which-way information, we must conclude that this is not a quantum eraser. However, the fact is, we are unable to determine the which-way information, not even in principle, because the labeling of the photons by the BBO crystal is a purely quantum process over which we have no control.

So, an ambiguity arises, which should make it clear again how misleading it is to think in terms of single photons traveling on deterministic paths. The supposed interrelation between interference patterns and the which-way information should always be taken with a grain of salt.

In the end, this doesn't look like a particularly interesting experiment: Whatever detector we choose to look at, we see only boring clumps of normal distribution patterns that show no signs of interference fringes.

The interesting part, however, comes from the coincidence counts—that is, the joint detection rates between detector D_0 and detector D_1 or detector D_2. Say we count all the events in which detectors D_1 and D_0 clicked jointly. That is, while the step motor provides for a linear spatial displacement of detector D_0 scanning the interference pattern of the signal photons, one keeps only those events in which detector D_1 also clicks, revealing the idler photon. By doing so, one obtains the joint detection rate between detector D_0 and detector D_1, R_{01}, interference fringes appear, like in Fig. 24 left.

While, looking at the joint detection rates between detector D_0 and detector D_2, R_{02}, interference phenomena are again manifest, but in form of anti-fringes, like that of of Fig. 24 right. Equivalently, one could say that the graph of R_{02} is a π-phase shift of the graph of R_{01}.

Because it has been a source of much confusion, it can't be emphasized enough that (similar to what we saw in the SWE and Walborn experiments), these are not the representations of the interference patterns one would measure at some detector. Instead, what Fig. 23 and Fig. 24 show is the correlation between a coincidence count of detector D_0 with that of detectors D_3, D_1, and D_2, respectively.

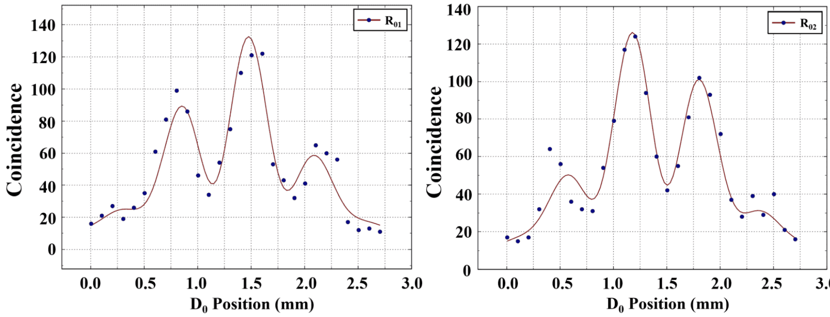


Fig. 24 Joint detection rates R_{01} and R_{02} (Kim et al. [9].)

They are jointly detected subsets of the signal photons of D_0 filtered out from a superimposed data stream. The interference pattern registered at D_0 can be extracted from the bell-shaped curve only after the idler photons have triggered $D_{1\text{-}4}$. They can be seen only retroactively. It is not possible to deduce what will happen to the idler-photons by observing the signal photons alone. To put it in other words, one 'selects' and 'picks out' only the data of the clumpy pattern without fringes where two detectors 'clicked' almost simultaneously—that is, those pixels activated by the signal photon at D_0 that correlate with an event triggered by the idler photon at D_1, D_2, or D_3. However, as already pointed out, a detection screen in front of these detectors would not show any interference fringes—not at detector D_0 nor at any of the other ones. They would all display the normal distribution. The sum of the data $R_{01}+R_{02}$ of the left and right graphs of Fig. 24 represents the real pattern measured by detector D_0, which would then appear similar to the curve of Fig. 23.

However, the that fact is, because these correlations are made, interference fringes and anti-fringes appear and every time a signal photon displaces itself on a fringe at detector D_0, then detector D_1 clicks. On the other hand, if it 'runs into' an anti-fringe, detector D_2 clicks. The apparently mysterious and weird thing about all this is that the question arises: How does the idler photon know whether it has to trigger detector D1 or detector D2 considering that the signal photon and detector D_0 are too far away (space-like separated) to convey any information? Once the signal photon collapses onto detector D_0, say, onto a fringe, it has no time to 'inform' the idler photon *"I hit the fringe, please trigger detector D_1"*. However, nevertheless, facts show that the idler photon indeed does 'know' whether the signal photon was a fringe or anti-fringe one and, thus, behaves accordingly.

At first, we might speculate that some form of FTL action from detector D_0 to the idler photon is at work here. Or, as many did, that this implies a temporal quantum retro-causation must be invoked to resolve the paradox.

According to this latter idea, the process goes the other way around: The idler photon, once it has randomly triggered either detector D1 or detector D2, eventually also, long after the signal photon triggered D_0, acts back into the past, telling the signal photon what to do. Or, if you prefer, the signal photon receives, from the future, information about whether the idler photon triggered detector D1 or detector D2 and, therefore, behaves accordingly, hitting the fringe or anti-fringe along the scan of detector D_0. The delayed choice seems to change the outcome of an event in the past. Effects seem to precede the cause, changing the order of the causal sequence.

This has caused a plethora of speculations and discussions that persist and continue to be spread all over the Internet. You will find YouTube videos and lots of discussions on forums, blogs, and social media, falsely claiming that through this experiment (and the others we aforementioned), QM has supposedly shown the existence of retro-causal action. While there are, indeed, good reasons to believe that the physics we know of does not necessarily rule out retro-causation (see, later, the chapter on the time-symmetric interpretation of QM), we are going to show that there is no need to resort to 'back-to-the-future' or 'time-machine' narratives to explain this experiment inside the conventional temporal order of the cause-and-effect paradigm.

First, let us not forget that we are dealing with a quantum system of two photons which are in a superposition state (both the signal and idler photons emerge from both slits) and, at the same time, are entangled (with anti-correlated orthogonal polarization states). There is an interplay and simultaneity of quantum entanglement and superposition—something with which we are already familiar due to the experiment of Walborn et al.

Second, the whole phenomenon can become meaningful only when we accept that the state collapse of the signal photon at detector D_0 does ***not*** cause a complete collapse at the idler photon which will remain in a superposition state. The state reduction of the signal photon to a particle state (a dot, a pixel on a screen, or a 'click' of detector D_0) 'removes' the entanglement, but does not cause state reduction to a particle state of the once-entangled idler photon, which remains in a superposition state. (In more technical terms, the measurement at D_0 reduces the system from a 'pure state' to a 'mixed state'; more on this in chapter IV 5). If the signal photon is absorbed, the idler is still 'in-flight' in a superposition state. Eventually, what is going on is quantum steering.

To clarify this formally in detail let us first use the following nomenclature: $|P_\alpha\rangle_\mathrm{S}$ is the ket-quantum state in Dirac notation with polarization P= H, V, $\pm 45°$ with α = s, i standing for the signal or idler photon and S= A, B for slits A or B. For example, $|H_s\rangle_\mathrm{A}$ is the state vector for the signal photon with horizontal polarization and emerging from slit A;

$|V_i\rangle_B$ stands for the idler photon with vertical polarization from slit B, and so on.

With this convention let us express how the signal photon must be in an orthogonal polarization superposition state ($|V_s\rangle_A + |H_s\rangle_B$ or $|H_s\rangle_A + |V_s\rangle_B$) and also be entangled with the idler photon which must be in the opposite (anti-correlated) superposition ($|H_i\rangle_A + |V_i\rangle_B$ or $|V_i\rangle_A + |H_i\rangle_B$, respectively). Therefore, the overall state function before detection at D_0 is:

$$|\Psi\rangle = \frac{|H_i\rangle_A|V_s\rangle_A + |V_i\rangle_B|H_s\rangle_B}{\sqrt{2}} + \frac{|V_i\rangle_A|H_s\rangle_A + |H_i\rangle_B|V_s\rangle_B}{\sqrt{2}}. \quad \textit{Eq. 9}$$

Before writing down the final process which is physically taking place, let us adopt a pedagogical bottom-up approach which first clarifies the different 'pieces' of the overall picture with which we are dealing. To fix the ideas, and by keeping in mind the anti-correlation of type-II entangled photons, consider, for example, that when the signal photon collapses at detector D_0 to the vertical polarization as coming from slit A (state $|V_s\rangle_A$), then the idler photon, if it is traveling towards the other detectors $D_{1\text{-}4}$ must still be in superposition as coming from the same slit A but with the opposite horizontal polarization state (state $|H_i\rangle_A$), while that from slit B must be in the vertical polarization state (state $|V_i\rangle_B$). The same applies when the signal photon collapses to the horizontal polarization state at slit B (state $|H_s\rangle_B$). That is, the quantum state of the idler photon after detection of the signal photon at D_0, but before being detected at $D_{1\text{-}4}$, is in shorthand:

Measurement at D_0: $|V_s\rangle_A$ or $|H_s\rangle_B$

↓

State of idler photon: $|\Psi_i\rangle_1 = \frac{|H_i\rangle_A + |V_i\rangle_B}{\sqrt{2}}$

Similarly, for the other two possible outcomes.

Measurement at D_0: $|H_s\rangle_A$ or $|V_s\rangle_B$

↓

State of idler photon: $|\Psi_i\rangle_2 = \frac{|V_i\rangle_A + |H_i\rangle_B}{\sqrt{2}}$

If you see that, we can refine this picture and note that the signal photons detected at D_0 as $|V_s\rangle_A$ or $|H_s\rangle_B$ must have been the result of the collapse of the signal photon superposition state

$$|\Psi_s\rangle_1 = \frac{|V_s\rangle_A + |H_s\rangle_B}{\sqrt{2}},$$

while $|H_s\rangle_A$ or $|V_s\rangle_B$ that of

$$|\Psi_s\rangle_2 = \frac{|H_s\rangle_A + |V_s\rangle_B}{\sqrt{2}}.$$

Because of what we discussed in chapter I.2 (especially with Eq. 1), the polarizations are orthogonal, $|\Psi_s\rangle_1$ and $|\Psi_s\rangle_2$ as also $|\Psi_i\rangle_1$ and $|\Psi_i\rangle_2$, represent the clumpy refraction pattern without any interference fringes. However, we also saw that, in the diagonal basis given by Eq. 3 and Eq. 4, these split up into symmetric and anti-symmetric components, as described by Eq. *5* and illustrated in Fig. 11.

Therefore, we can summarize the two cases which show the photon's state in the rectilinear and diagonal base with the following diagrams throughout the next two pages. These illustrate the process that leads, at the end of the line, to the observed detector responses.

Case I: detector D_0 measures $|V_s\rangle_A$ or $|H_s\rangle_B$.

Before measurement at D_0:

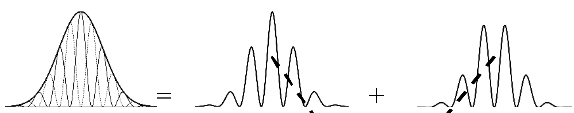
$$|\Psi\rangle = \frac{|H_i\rangle_A|V_S\rangle_A + |V_i\rangle_B|H_S\rangle_B}{\sqrt{2}} + \frac{|V_i\rangle_A|H_S\rangle_A + |H_i\rangle_B|V_S\rangle_B}{\sqrt{2}}$$

Collapse at D_0

$|V_s\rangle_A$ or $|H_s\rangle_B$.

These originated from the signal photon state vectors:

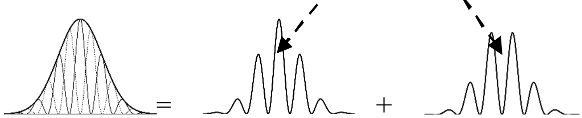
$$|\Psi_s\rangle_1 = \frac{|V_S\rangle_A + |H_S\rangle_B}{\sqrt{2}} = \overset{(a)}{\frac{|45°_s\rangle_A + |45°_s\rangle_B}{\sqrt{2}}} + \overset{(b)}{\frac{-|-45°_s\rangle_A + |-45°_s\rangle_B}{\sqrt{2}}} \qquad \textit{Eq. 10}$$



$$|\Psi_s\rangle_1 = |\Psi_s\rangle_1^{\text{Sym}} + |\Psi_s\rangle_1^{\text{Asym}}$$

Measurement at D_0 projects also the ('in-flight') anti-correlated idler photon into:

$$|\Psi_i\rangle_1 = \frac{|H_i\rangle_A + |V_i\rangle_B}{\sqrt{2}} = \overset{(a)}{\frac{|45°_i\rangle_A + |45°_i\rangle_B}{\sqrt{2}}} + \overset{(b)}{\frac{|-45°_i\rangle_A - |-45°_i\rangle_B}{\sqrt{2}}} \qquad \textit{Eq. 11}$$



$$|\Psi_i\rangle_1 = |\Psi_i\rangle_1^{\text{Sym}} + |\Psi_i\rangle_1^{\text{Asym}}$$

Collapse at D_1:	Anti-fringe	Fringe
Collapse at D_2:	Fringe	Anti-fringe

Diagram 1

The difference in the signature of the anti-symmetric component (between Eq. 10b and Eq. 11b or between Eq. 12b and Eq. 13b) is a mere algebraic one which appears by exchanging the order of polarizations or slits (Convince yourself by inserting Eq. 3 and Eq. 4 into $|\Psi_s\rangle_{1/2}$ and $|\Psi_\mathrm{i}\rangle_{1/2}$). You may now realize the meaning of the dashed arrows. As an example, consider case I. If detector D_0 detects one signal photon in a fringe position, that is, collapses to symmetric state vector $|\Psi_s\rangle_1^{\mathrm{Sym}}$ (Eq. 10a), it must project the anti-correlated idler photon into the 'anti-fringe state', that is, to $|\Psi_\mathrm{i}\rangle_1^{\mathrm{Asym}}$ (Eq. 11b).

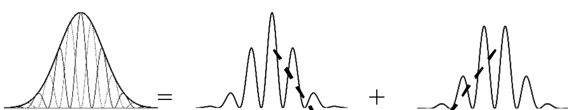
Case II: detector D_0 measures $|H_s\rangle_A$ or $|V_s\rangle_B$.
Before measurement at D_0 :

$$|\Psi\rangle = \frac{|H_i\rangle_A|V_s\rangle_A + |V_i\rangle_B|H_s\rangle_B}{\sqrt{2}} + \frac{|V_i\rangle_A|H_s\rangle_A + |H_i\rangle_B|V_s\rangle_B}{\sqrt{2}}$$

Collapse at D_0 ↓ $|H_s\rangle_A$ or $|V_s\rangle_B$.

These originated from the signal photon state vectors:

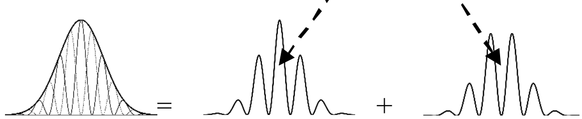
$$|\Psi_s\rangle_2 = \frac{|H_s\rangle_A + |V_s\rangle_B}{\sqrt{2}} = \overset{(a)}{\frac{|45°_s\rangle_A + |45°_s\rangle_B}{\sqrt{2}}} + \overset{(b)}{\frac{|-45°_s\rangle_A - |-45°_s\rangle_B}{\sqrt{2}}} \quad \textit{Eq. 12}$$



$$|\Psi_s\rangle_2 = |\Psi_s\rangle_2^{\mathrm{Sym}} + |\Psi_s\rangle_2^{\mathrm{Asym}}$$

Measurement at D_0 projects also the ('in-flight') anti-correlated idler photon into:

$$|\Psi_i\rangle_2 = \frac{|V_i\rangle_A + |H_i\rangle_B}{\sqrt{2}} = \overset{(a)}{\frac{|45°_i\rangle_A + |45°_i\rangle_B}{\sqrt{2}}} + \overset{(b)}{\frac{-|-45°_i\rangle_A + |-45°_i\rangle_B}{\sqrt{2}}} \quad \textit{Eq. 13}$$



$$|\Psi_\mathrm{i}\rangle_2 = |\Psi_\mathrm{i}\rangle_2^{\mathrm{Sym}} + |\Psi_\mathrm{i}\rangle_2^{\mathrm{Asym}}$$

Collapse at D_1:	Anti-fringe	Fringe
Collapse at D_2:	Fringe	Anti-fringe

Diagram 2

Whereas, if detector D_0 detects the signal photon at an anti-fringe position, that is, collapses to anti-symmetric state vector $|\Psi_s\rangle_1^{\mathrm{Asym}}$ (Eq. 10b), it must project the anti-correlated idler photon into the 'fringe-state', $|\Psi_{\mathrm{i}}\rangle_1^{\mathrm{Sym}}$ (Eq. 11a). Exactly the same sort of correlations applies to case II.

Of course, which of the two cases will happen each time is again a completely quantum random process. That's why, after several fringe and anti-fringe signal photons accumulate at D_0 in a normal distribution curve, leaving no trace of interference fringes. The very same process takes place for the idler photons which will display the bell-shaped curve at the other detectors as the sum of the symmetric and anti-symmetric interference patterns as well.

At this point, the decisive insight that becomes clear from all this is that, due to quantum steering, immediately after the measurement of the signal photon at detector D_0, but before the still 'in flight' idler photon is measured at detectors $D_{1\text{-}4}$, it carries this information. At this temporal stage, the idler photon already 'knows' whether its signal partner was projected onto a fringe or anti-fringe and will behave accordingly. No quantum retro-causal effect must be invoked. The measurement at D_0 of the signal photon already determines, a priori, the probabilities that the idler-photon will hit either $D_{1\text{-}4}$. The signal photon appears to be 'clairvoyant' only if we overlook this step of the process and forget to look at things from the perspective of the diagonal eigenbasis.

Still, this does not completely resolve the apparent issue with a supposed retro-causal action of the idler photon when it makes detector D_1 or detector D_2 click. The attentive reader, who has thought this all through, might have realized that the mystery contains another piece requiring an answer. This is because, even if the idler photon already carries in its state vector the information about whether the signal photon hit a fringe or anti-fringe in D_0, how does beamsplitter BS 'know' whether it must direct the idler photon towards detector D_1 or detector D_2? After all, a beamsplitter is simply a piece of glass or a crystal that is not supposed to be a 'receptionist' that forwards messages according to its informational content. Rather, it is a transparent medium that splits a photonic stream with a prefixed probability (usually 50%) towards two (usually perpendicular) directions. There is no reason to believe that it will treat photons in a symmetric quantum state differently from those in an anti-symmetric state.

Therefore, the next step is to clarify how the idler photon which, after the measurement in D_0 is no longer entangled with the signal photon, but is still 'in-flight' in one of the four possible quantum superposition states $|\Psi_{\mathrm{i}}\rangle_1^{\mathrm{Sym}}$, $|\Psi_{\mathrm{i}}\rangle_1^{\mathrm{Asym}}$, $|\Psi_{\mathrm{i}}\rangle_2^{\mathrm{Sym}}$ or $|\Psi_{\mathrm{i}}\rangle_2^{\mathrm{Asym}}$, will reach detector D_1 or D_2. What kind of signal will it trigger?

First of all do not allow your mind to fall into the separation and measurement fallacy and don't forget that there is only ***one*** photon taking ***two*** paths A and B, being reflected at ***both*** mirrors M_A and M_B and being reflected ***and*** transmitted at beamsplitter BS. Second, recognize that, despite a perfect physical symmetry represented by the part of the experimental setup made of the mirrors M_A and M_B, beamsplitter BS, and detectors D_1 and D_2, an optical anti-symmetry nevertheless holds: This part of the system behaves differently according to the photon's diagonal polarization being in a $+45°$ or $-45°$ superposition state. It can be shown (for a detailed discussion, see Appendix 0) that, if a measurement at D_0 selects the symmetric part of the signal photon ($|\Psi_s\rangle_1^{\text{Sym}}$ or $|\Psi_s\rangle_2^{\text{Sym}}$), projecting the idler photon into the anti-symmetric state ($|\Psi_i\rangle_1^{\text{Asym}}$ or $|\Psi_i\rangle_2^{\text{Asym}}$), once reflected at ***both*** mirrors M_A and M_B, and after being reflected ***and*** transmitted at beamsplitter BS, the so transformed idler photon's state function will displace it on a fringe at D_1 and on an anti-fringe at D_2. Vice versa, if a measurement at D_0 selects the anti-symmetric part of the signal photon ($|\Psi_s\rangle_1^{\text{Asym}}$ or $|\Psi_s\rangle_2^{\text{Asym}}$), projecting the idler photon into the symmetric state ($|\Psi_i\rangle_1^{\text{Sym}}$ or $|\Psi_i\rangle_2^{\text{Sym}}$), once reflected at ***both*** mirrors M_A and M_B, and after being reflected ***and*** transmitted at beamsplitter BS, the so transformed idler photon's state function will be shifted by 90° compared to that of the previous case and will displace it on an anti-fringe at D_1 and on a fringe at D_2.

To cut a long story short: Detectors D_1 and D_2 will always react in a complementary fashion and in accordance with the signal photon's symmetry state. The two cases switch permanently and randomly for each photon, as described above, and the overall result is nevertheless the normal distribution. However, if one considers how detector D_0 takes the measurements of the signal photons guided by a step motor along the perpendicular direction of its propagation, this explains why the joint detection rate between detectors D_0 and D_1, R_{01}, displays a standard Young's double slit interference pattern, while the joint detection rate between detectors D_0 and D_2, R_{02}, displays the complementary π-phase shifted interference pattern of Fig. 24.

Finally, also in this case, just as with what we have seen with in the experiment with the polarizers of Fig. 10, the SWE and Walborn et al. experiments, one can again explain everything inside an orthodox cause and effect paradigm without any resorting to retro-causality. Once we become aware of the separation and measurement fallacy, avoid retro-ductive reasonings of counterfactual definiteness, and keep the potential quantum steering effects in mind, then the retro-causal hypothesis appears in all its deceitfulness. On the other hand, these sorts of experiments reinforce our deeper understanding of the foundations of QP and show how ‘quantum ubiquity’ is at work. It is not just a mathematical figment, rather, must have an ‘inherent element of physical reality’, as ‘someone’ used to say, even though he was looking at it from the opposite standpoint.

II. Interpretations of Quantum Mechanics

You will have realized that while QP is a rigorous and exact science, it has no clear accepted ontology. Formally, everything is consistent but it relies on no particular models of reality. Its power resides precisely in the fact that it is a mathematically complete theory that can predict the outcome of experiments with high precision, successfully describing all the quantum processes of the real world. After all, this is what science is primarily about. Models, interpretations, ontological questions, and philosophical speculations are sometimes added but only as tolerated addenda, not as necessary ingredients of a scientific theory. This quickly led physicists to realize that it is easier to restrict oneself to the calculations and to follow the '*shut up and calculate*' approach without bothering much about the ontological model of the quantum world. And because the mathematics involved in learning and using consistently modern QT takes a huge amount of time, most physicists do not allow themselves to go beyond that.

However, as we have discussed previously, some famous physicists were an exception to this approach. For example, while Einstein wasn't satisfied with mere calculations, he also didn't try to construct a new worldview emerging from the quantum phenomena. Rather, he simply hoped that some form of deterministic local realism could be saved. As you know, this attempt ended in failure and, particularly because it became clear that QM violates Bell's inequalities, today almost no one follows Einstein's path. Most physicists (still) stick to the Copenhagen interpretation of QM, which we already discussed and that, in any case, doesn't appear to be an 'interpretation' at all but, rather, simply a working attitude that discourages speculation beyond the empirical facts and its rigorous formalization. For this reason we have not considered the Copenhagen interpretation here, we already discussed it in Vol. I.

Other physicists, philosophers, and mathematicians have tried hard to build models of reality which are compatible with the current structure of QM. In the following sections, we will take a look at few of these interpretations but the description will necessarily be somewhat superficial and incomplete, as it is impossible to do justice to each of them with only a few pages. Nowadays, one can count about two dozen different interpretations of QM, and it would be impossible to illustrate them all in detail. But precisley this proliferation of interpretations demonstrates the controversial nature of this issue. Additionally, you will by no means find a solution to the quarrel here. How the weird world of QP is to be correctly interpreted remains a widely debated and unresolved question and is still more a matter of personal taste and preference than a real scientific matter. The aim of this section is therefore only to give you an intuitive glimpse into a few mainstream interpretations of QP out of many, so that you will be able to proceed further and judge for yourself whether you believe that one or the other theory merits more or less attention.

With that said, at the end of this chapter the author feels authorized to present a somewhat biased concluding remark to clarify why he isn't at all passionate about what are called 'interpretations' of QM. I added this chapter only out of completeness but feel that physics is not going in the right direction by insisting so much on creating such a profusion of diverse interpretations.

1. The de Broglie-Bohm pilot wave interpretation

The '*de Broglie-Bohm interpretation*', also called the '*Bohm pilot-wave theory*' or '*Bohmian mechanics*' (BM), is probably the most notorious and, in a certain sense, most surprising interpretation of QM. It was proposed by L. de Broglie way back in 1927. He presented it at a conference attended by W. Pauli, who pointed out some apparent inconsistencies to which de Broglie could not reply in a satisfying manner. The story goes that de Broglie perceived this as a publicly humiliating experience and soon gave up further attempts to develop the theory. In 1952, however, D. Bohm rediscovered and took up de Broglie's idea, showing how Pauli's objections were unwarranted. An interesting historical anecdote that could be a lesson, especially for young researchers: Never be intimidated by authority!

The key characteristic of this approach to QM is that BM can recover classical determinism, though at the cost of local realism. QM must be at least a non-local theory or a non-deterministic theory or both, as the violation of Bell's inequality and its related theorem clearly tell us. In fact, Bell's theorem does not necessarily give away determinism and hidden variables. A non-local but deterministic theory is still allowed. This is precisely what Bohm strived for, successfully. Bell *"saw the impossible done,"* as he commented, referring to BM, and later embraced the reality model that it suggests.

In BM, one still conceives of classical particles, that is, point-like entities as our intuition suggests, as little pinpointed physical objects (with or without an inner structure) that have a definite actual position following a deterministic trajectory. This aspect makes the approach so appealing to many physicists, who are uneasy about giving up a classical Laplacian hidden variable determinism.

But how can particles exhibit the interference phenomenon that QP so ubiquitously displays? De Broglie's groundbreaking idea was to separate the particles from the wavefunction as two different categories of reality. The former has precise initial conditions (position and velocities in a 3D space, the so called *'phase space'*) while the latter is no longer only a mathematical statistical tool to describe probabilities but, rather, a fundamental object with its own ontology: It is the '*pilot wave*' that 'guides' the particles along deterministic paths. It is this background wave that determines the particles' dynamics in all the Universe. BM is still based on Schrödinger's equation but the wavefunction with which it works has the status of a sort of fluid with stochastic fluctuations that drive the process from an initial non-equilibrium to a quantum equilibrium state and which determines all the particles' positions, the '*configuration space*.' The hidden variables are the unknown random initial positions of the particles which convey to QM its randomness and unpredictability.

The manner in which this pilot wave acts can be seen in the example of the two-slit experiment in Fig. 25. The particles travel through one or the other slit, according to their initial conditions, but the 'guiding wave' forces them to follow trajectories that the interference pattern dictates. It is the wave that guides the particles that interferes, and not the particles themselves.

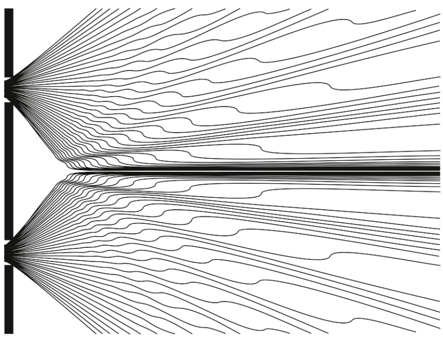


Fig. 25 The particles' trajectories according to BM.

This implies that some sort of force must act on the particles and distribute them along specific paths, avoiding the black fringes but still exhibiting the wavy nature of the process: a force described by a '*quantum potential*.' This allows the pilot wave to displace the particles only at the positions corresponding to the interference fringes, mimicking the wave-particle duality which, in the context of BM, appears to be no duality at all.

Introducing this quantum potential, the Schrödinger equation appears as a 'guiding equation.' Bohm imagined particles as having an inner structure through which this universal quantum potential somehow drives all the particles of the Universe as a whole: the wavefunction of the Universe.

How fluid dynamics can mimic QM was more recently illustrated with real experimental water fluid-dynamic systems that, indeed, can reproduce surprisingly well the quantum-mechanical phenomena, such as the two-slit experiment. A drop of fluid, such as an oil drop, striking the surface of a fluid bath produces waves that, in turn, propel the droplet across the bath. [10] This can't be mere coincidence, can it? This is still unclear, because more refined experiments of fluid mechanics have partially falsified this droplet model. Other researchers failed to obtain the same results. [11] Whether this is due to the fact that these experiments are based on purely classical mechanical principles that cannot take into account non-local correlations, or simply because BM is wrong, remains to be seen.

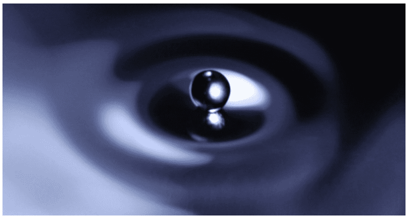


Fig. 26 Fluid dynamical simulation of BM. MIT News - Image: Dan Harris.

It is interesting to note, however, that in BM, unlike the property of position, the other properties of particles, such as mass, spin, charge, etc., are not definite and localized at the position of the particle but, rather, spread out over all the wavefunction. The non-local character emerges from the interconnectedness of all the particles with each other: The momentum and acceleration of all the particles of the Universe are dependent on each other simultaneously. With this context in mind, Bohm developed a holistic view of the Universe, which he divided into an *'implicate order'* and an *'explicate order'*. The implicate order connects everything with everything else, where any individual element could reveal *"detailed information about every other element in the universe"* in an *"unbroken wholeness of the totality of existence as an undivided flowing movement without borders."* [12] Meanwhile, the explicate order is the manifest world. It is secondary and derivative and it flows out from the laws of the implicate order.

Supporters of BM frequently point out how it accounts successfully for all non-relativistic quantum physics. Modern relativistic extensions seem to be possible as well, although these are not yet as fully developed as their non-relativistic counterpart. From this perspective, BM mechanics seems to shed light on many of the mysterious aspects of QT. For example, there is no 'quantum collapse' as in classical QP. The collapse exists only if we conceive of the wave-particle duality as an intrinsic property of particles. Meanwhile, in BM, particles always remain as such and, independently, the pilot wave remains a wave. Heisenberg's uncertainty principle arises as a natural consequence of this guiding wave. (See, for example, our discussion in Vol. I of the uncertainty principle derived from the diffraction and interference with the pinhole experiment.) In the experiment testing Bell's inequality violation, the action at a distance between the two polarizers is not something occurring between the particles but, rather, something that happens between the pilot waves associated with the two particles. In this sense, local realism for each particle is safe and the non-local character of QT is ascribed to the pilot wave alone. Moreover, while in classical QM the Born rule is not that obvious (why is a probability density given by the modulus squared of the wavefunction and not by a modulus raised to any other even power?), Bohm was able to show how the Born rule emerges naturally once a quantum system reaches the state of quantum equilibrium.

Overall, BM is a quite fascinating interpretation and alternative to the standard Copenhagen interpretation and mindset. It is, to date, the model of microphysical reality that comes closest to our macroscopic intuition.

2. The Many Worlds Interpretation

The '*Many Worlds Interpretation'* (MWI) or '*relative state interpretation*' of QM is another attempt which, along with BM, finds much attention among philosophers and physicists. Hugh Everett introduced it in 1957 with his Ph.D. thesis entitled "*The Theory of the Universal Wavefunction*." (Wheeler was his advisor.) It furnishes, first and foremost, a possible answer to the measurement problem. Its basic idea is that what we observe to be the collapse of the wavefunction, as Nature's mysterious selection of only one possible realization of the state of a system out of many, eventually infinite possible outcomes, can easily be explained away if we admit that we are living in a Universe that, at every instant, branches into many 'worlds,' each of which realizes just one possible state of the measuring device or some other macroscopic variable. For example, the double-slit experiment splits our actual universe into many timelines according to the granularity of the photographic screen (however, not according to all the possible outcomes for which the wavefunction allows, as in the sum-over-histories Feynman path integrals).

By 'Universe,' we mean the collection of all the possible worlds. (This type of 'multiverse' as a collection of many Universes should not be confused with the modern Multiverse theories so in fashion nowadays.) For instance, if QM tells us that there is a 50% chance that the measurement of an electron's spin will furnish a spin-up or spin-down result and that we measure spin-up, in the MWI this means that at the instant of measurement, our present world branches into two parallel ones: one in which we (or the measurement apparatus) find ourselves with the spin-up electron and another in which there is a 'copy' of ourselves (or an identical copy of that measurement apparatus), but with the spin-down electron. Every time another measurement is performed, each of these worlds again splits into two worlds (see Fig. 27), giving rise to a process whereby the number of timelines grows exponentially. This is supposed to happen not only for a measure of a human observer but for each quantum process at each moment.

Therefore, in the MWI, what did not manifest in the past in this world did manifest in some other. In its entirety, the whole Universe realizes all the possible 'histories' for which QP allows.

This permanent 'splitting' of ourselves into gazillions of universes parallel to our own every time a quantum mechanical event occurs may sound like a quite extravagant hypothesis that is difficult to accept. However, it does explain several weird quantum phenomena and is taken seriously by some physicists.

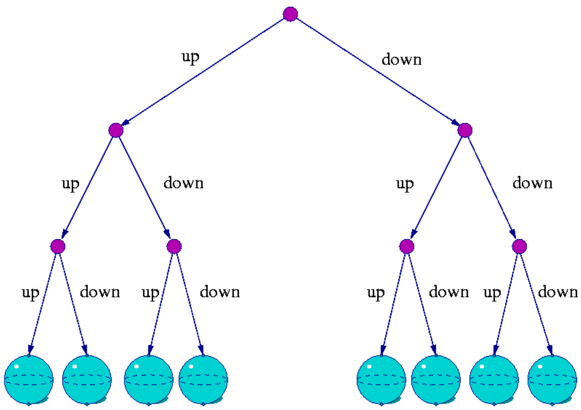


Fig. 27 Spin-up and spin-down in the MWI

For instance, the source of quantum randomness in this perspective appears to be no randomness at all, as all the possible outcomes come truly into existence. However, we become aware of only one of these outcomes because we find ourselves evolving along the only one of the infinite branches of these myriad worlds where just that result is realized. Deterministic paths do not exist in the single 'world' but, rather, in the Universe as a whole.

Quantum superposition also appears to no longer be so mysterious: Schrödinger's cat is alive in one of these worlds, while it is dead in another parallel world. There is no real superposition of states, only the actualization of both states but each realized in different realities. And, contrary to Bohm's theory, even locality is recovered.

In fact, entanglement and its non-local character appear only to the observers inside the single world but have no need to exist in the Universe as a whole plurality of these worlds, as the correlations or anti-correlations in the Bell experiments exist in the individual worlds, not in the Universe.

Quantum decoherence can be interpreted as a branching phenomenon along one tree of the Universe. The collapse is not triggered by the measurement apparatus (and even less by an 'observer') but, rather, is something which occurs continuously, as each actual world splits into all the many worlds that the wavefunction predicts at any instant, with the probability representing the relative amount of similarity between one and the other.

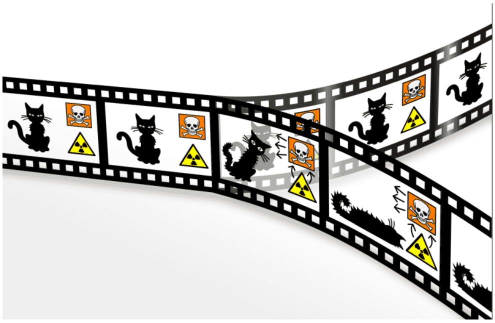


Fig. 28 Schrödinger's cat paradox solved in the MWI.

Quantum superposition amounts to the coexistence of many worlds which, however, cannot communicate or have any physical contact with each other.

The reason why such an interpretation meets the sympathy of high-ranking physicists is that, apart from the annoying multiplication of worlds, it satisfies the desire to maintain things as far as possible in the frame of an at least partially Newtonian mechanical worldview of cause and effect describing reality in terms of particles, actual trajectories, and classical deterministic causes and effects.

3. Superdeterminism

In Vol. I, we cited Bell's description of, and comment about, this approach to QM which, in principle, explains the violation of his inequalities and reintroduces through the back door, so to speak, a hidden variable local realistic theory. '*Superdeterminism'* is, as the name suggests, an even stronger form of determinism that Bell conjectured to explain the correlation between entangled particles measured by the two distant spacelike separated polarizers without resorting to non-locality. Superdeterminists point out that, even in a conventional deterministic view, the choice of the experimenter, such as his/her choice of how to orient the polarizers, remains free. This free choice is still an indeterministic aspect of the classical deterministic conception that experimenters have in mind. (After all, nobody likes to think that his/her own choices are predetermined and fixed.) Superdeterminism imposes a further restriction by considering this choice as being predetermined as well, in the sense that there is no such thing as free will. Under this approach, free will may be only an illusion and, in fact, is determined only by the laws of physics. For example, the choice that the experimenter makes may have been predetermined when the particles were produced in the source and somehow already 'knew' what the experimenter would do once they reached the distant polarizers. Eventually, one can go so far and believe that every single event—and, therefore, the most minute behavior of complex living beings—was predetermined much in advance at the time of the Universe's creation. This information might be somehow retrievable by the entangled particles which already "know" the polarizer orientations they will find and behave accordingly. The experimenter's free choice is an illusion because it forms part of the initial conditions which determine the outcome of an experiment later. An extremely fine-tuned Universe, indeed. As strange and extreme as this might sound, it is still a logical possibility that some physicists (such as Nobel laureate Gerard 't Hooft) and a few philosophers of science take seriously because it eliminates the necessity for 'spooky actions at a distance,' and determinism and local realism could be vindicated.

4. Objective collapse theory

Most physicists like to conceive of reality in terms of particles governed by deterministic dynamics of causes and effects, dismissing the wavefunction as being not a real entity and (more or less unconsciously) aiming to recover a local deterministic realism, ideally à la Laplace, with no place for inherently cause-less random phenomena. There are, however, also some who take the middle ground, retaining the particle-picture of classical mechanics but admitting pure randomness as an unavoidable aspect of the microphysical realty. This is the case in the '*objective collapse theory'* (OCT), also called the '*spontaneous localization model*' or '*dynamical reduction program*.'

This interpretation, first proposed in 1986 by Giancarlo Ghirardi, Alberto Rimini, and Tullio Weber (and therefore also known as the GRW theory), aimed to explain and make sense of the measurement problem. It did not introduce quantum potentials and objectively real universal wavefunctions suggested by BM or postulate a plethora of branching worlds like in the MWI. Instead, it conjectures that the wave-packet reduction is a real physical process. Here, the quantum state reduction (or projection of the state vector onto one of its basis eigenstates) is described by a spontaneous process of localization which occurs instantly at a micro-physical level. Superposition of states exists but is quickly suppressed by some (unspecified) trigger mechanism that causes particles to localize in space. It is not the act of measurement that causes state reductions or 'collapses' of the wavefunction, as in the MWI and in BM. Rather, these occur continuously on a microscopic scale. According to the OCT, there is, in the Universe, a spontaneous localization of the particles which is continuously happening at random times. It is a sort of background noise of space, something reminiscent of *'Brownian motion'* that molecules display due to thermal excitation (from botanist Robert Brown who observed in 1827 under a microscope grains of pollen zigzagging in water, motion that was later best explained in a complete physical theory by Einstein, and for which Einstein received his Nobel prize), and it causes the state vectors to be projected to one of its possible eigenstates. This objective localization, or objective reduction, determines the fuzziness of the spatial distribution of particles, the coming and going of virtual particles, as any other random quantum phenomenon. Moreover, the theory predicts that this spontaneous localization mechanism is extremely rare for the single particle, but becomes more efficient with a growing number of particles. This amplification of the spontaneous reduction in a quantum system of many particles explains why no quantum effects are observed in a macroscopic world. Recall how decoherence is only one of the ingredients necessary to solve the Schrödinger’s cat paradox; OCT completes this resolution by postulating a spontaneous reduction mechanism. The OCT is, therefore, a fundamentally indeterministic theory which conceives of point-like particles as emerging from a process of spontaneous localization. The OCT also calculates the frequency with which this ontologically real wavefunction collapse occurs, depending on the number and mass of the particles, thereby paving the way for possible future experimental verifications. This latter aspect makes OCT stand out among other interpretations.

5. Time symmetric quantum mechanics

Not only in QP, but more generally in all physics, time appears as a parameter (in fact, recall how time isn’t an observable in QM but is just a parameter) that does not distinguish the present state of a physical system from other, past or future states. While our everyday experience conceives of the present instant, which we call the ‘now,’ as the only real state of existence, and considers that time flows in only one direction, from the past to the future, there is nothing in physics that gives to the past and future a less important ontological status than that ascribed to the present. In physical theories, time is simply a number that can take on a value corresponding to past, present, or future events without preference. This fact, seemingly at odds with our perception and everyday experience of the world, has caused a lot of debate and discussion among physicists and philosophers. There is no clear understanding as to what time really is and why we perceive it as being unidirectional; nothing in physical theories suggests this. It is the famous problem of the '*arrow of time*' which, according to present understanding, finds a partial (though not full) solution in statistical physics.

To some physicists, this suggested that time-symmetry should be incorporated into QP. The '*time-symmetric interpretation'* (TSI) of QM, also called '*two-state vector formalism*,' was first proposed by the Japanese theoretical physicist Satosi Watanabe in 1955 and rediscovered by Yakir Aharonov in 1964. They realized how we force (more or less unconsciously) into physical theories our human prejudice, according to which we conceive of a system in the present moment as evolving only towards the future by setting some initial boundary conditions. From the outset, we structure our theories as a system evolving from the present to the future by imposing initial conditions in the present, ignoring a priori the possibility that physics also works well the other way around: The time-evolution of a physical system can be calculated equally well by imposing the final conditions instead.

In the TSI, one goes a step further by allowing time to flow in both directions: from a present state with initial boundary conditions towards the future, as also from a future state with final boundary conditions evolving towards the present. In conventional QM, one introduces a special evolution operator which acts on an initial pre-selected state represented by the state vector $|\Psi\rangle$ and 'evolves' it from a present time t_1 to a future instant t_2—that is, to a 'post-selected' state $|\Phi\rangle$. In the TSI one adds also the same evolution operator but acting onto a post-selected state vector $|\Phi\rangle$ evolving it from the future to the present time 'pre-selected' state $|\Psi\rangle$. This combination of the forward and the backward causation suggests, again, the possibility of retro-causation, though this appears in the theory only as a formal expedient, not a necessary reality. If the TSI implies quantum retro-causation, beyond a mere abstraction, remains a matter of debate. [13]

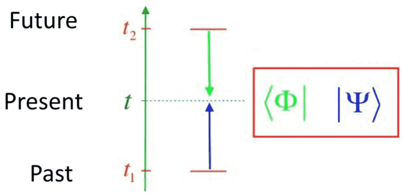


Fig. 29 The two-state vector interpretation of QM.

As a side note it might be worth mentioning how in this context, a '*weak measurement theory'* was developed—that is, a theory of quantum measurements that are weak enough to avoid significant interaction and disturbance but that nevertheless obtain information about a quantum system by averaging the measurements. (This, however, does not really avoid Heisenberg's uncertainty principle because the information is extracted from several measurements, not just a single particle.)

Unlike BM or the MWI, the aim of the TSI is not to re-interpret the typical quantum paradoxes but, rather, to furnish a different formal structure that takes time-symmetry seriously. It is, indeed, able to reformulate all the conventional QP without contradiction and, according to its founders, makes it, in part, even formally simpler. It is not a mainstream interpretation but the author has included it in this short list of alternatives because he is inclined to believe that it, together with the interpretation of QM in the next couple of sections, may indeed be onto something—perhaps a deeper truth which actually still escapes our understanding.

6. The transactional interpretation

Another similar but distinct approach from the TSI is the *'transactional interpretation'* (TI) of QM. Here, one deals with time in a peculiar and non-conventional manner as well. The TI of QM was first formulated in 1986 by J. G. Cramer, a professor at the University of Washington, who took up a theory of radiation proposed in 1945 by Wheeler and Feynman and further developed to a relativistic theory by R. E. Kastner. [14]

In its basic classical version, the TI conceives of a modified interaction theory between charged particles. The interaction between one particle and another particle is no longer seen as the emission of a single field propagation in space at a specific time from an emitting particle to an absorbing particle which will absorb a quantum (or the force mediating photon/boson) at a later time. Rather, it represents the emission/absorption as an exchange of two fields emanating from both particles at the same time and that meet halfway with a *'quantum handshaking'* reciprocal action. Both particles act with a temporal shift – the emitter with a future-directed half-retarded field emission, the *'offer wave'*, and the absorber with a past-directed half-advanced field response by means of a *'confirmation wave'*. This representation can equally describe a radiative process of energy transfer from an emitter to an absorber. It is called a 'transaction' because it is analogous to a financial and supply transaction.

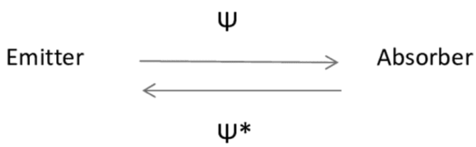


Fig. 30 The emitter/absorber process as an offer- and confirmation wave 'transaction'.

Such an interpretation also provides a natural explanation of Born's rule. (Did you notice? The MWI makes the same claim.) As we know, the probability density is given by the modulus square of the wavefunction $|\Psi|^2 = \Psi\Psi^*$, with the starred wavefunction being the complex conjugate. In the Schrödinger equation, the wavefunction Ψ represents a forward propagating wave, while its complex conjugation formally represents a backward propagating wave. This is a quite normal algebraic convention used throughout physics and engineering. (For example, see how we treated the polarized wave components $|E_x|^2$ and $|E_y|^2$ in Appendix A II.) The complex conjugation adds a little negative sign to the front of the time parameter of a wave but this is cancelled out once modulus squared and, usually, one does not attach to it a particular ontological significance, considering it something not physically real. Instead, in the TI, one takes this seriously, not just as a mathematical trick. The Born rule can be interpreted as an emitter/absorber interaction described by the product of an offer wave and a confirmation wave.

Moreover, in this picture, the collapse of the wavefunction is caused by the response of the absorber. The ontology of the measurement process (see the chapters in Vol. I on the wavefunction collapse, quantum decoherence, and the measurement problem), which has caused lots of debates and headaches, is seen as an irreversible process created by the absorber breaking the unitarity of the state evolution. (We will dwell further on this latter aspect in chapter IV.7.) The absorber's response is, in itself, viewed as the measurement.

7. Relational quantum mechanics

'Relational quantum mechanics' (RQM) is a theory that does not necessarily try to explain the paradoxes of QP either and that has a history and motivation which are somewhat different from those of the other interpretational attempts. It emerged from the need to find a theory of quantum gravity (QG). Modern approaches towards unifying the force of gravity with QM and GR have failed and many physicists agree that the main problem stems from our too trivial and naive understanding of notions like space, time, and space-time.

What modern QT and relativity still have in common, and what they have inherited from the Galilean and Newtonian mechanics, is that the concepts of space and time are posited as fundamental. While it is the case that relativity went a step further than CM by unifying these in a more plastic and relative notion as space-time where space and time can depend on the observer's frame of reference, its conceptual ontology has not really changed. Space, time, and space-time have not been questioned further because they present themselves to us as self-evident everyday experiential facts. After decades of research which has attempted to find a QG theory, it turns out that space-time might not be a fundamental property of the Universe but, instead, an emergent property.

After his research into '*loop quantum gravity*,' Carlo Rovelli, an Italian physicist also known to the public as an active popular science divulgator, introduced RQM in 1984. This theoretical framework seeks to reformulate QM without space and time as absolutes but, rather, as relative properties of the Universe, which otherwise have no meaning on their own. In relativity, it is already known that time passes by with different speeds for different observers, or that clocks tick faster in space than they do on Earth (as in any gravitational field). Moreover, relativity tells us that notions like simultaneity, speed, and a particle's position in space have no absolute value and meaning. Rather, they must be determined relative to a reference system. Similarly, in RQM, the state of a quantum system can be relative to the observer's state of motion or position or to other physical parameters characterizing it. Two different observers may see a quantum system in opposite states—for example, for one, it is in a superposition state whereas for the other, it is in an eigenstate. The very notion of 'state' becomes a relative notion that depends on the relation between the observed object and the measurement instrument. The wavefunction or state vector describes a correlation between two objects. It makes no sense to speak of a state of one or the other system if not in comparison to each other or to some reference object and quantum state. For example, it is meaningless to state that a physical process—say, the speed of light—is fast. One must always specify a reference process against which it can be measured. RQM is still in its infancy but seems to capture something essential that our subjectivity, based on an instinctive way of perceiving and understanding the world, seems to miss or unconsciously posit as granted.

8. On the proliferation of the interpretations of quantum mechanics

Lots of other interpretations were not mentioned here (such as the consistent histories, ensemble, modal interpretation, QBism, and many others). The author, however, does not feel it instructive to discuss each of them or to dwell further on those that have already been mentioned. So far, I have tried to focus mainly on the facts and experiments and, in this chapter, on laying out the different interpretations of QM, seeking to avoid being too biased by favoring a personal point of view. However, having delivered the basics and main concepts of the foundations of QP, I can finally engage in some gossip.

The simple fact that there is such a proliferation of quantum interpretations, most of which (with few exceptions) have no contact with each other, speaks volumes: There is actually zero consensus among physicists on how to interpret QP at all! Everyone sees something different and, of course, pretends that his/her interpretation or theory is that which best captures the meaning and/or ontology of quantum phenomena. There is a saying: 'There are as many interpretations of QM as there are physicists.' Perhaps there are even more than that.

Finally, whatever we believe, and despite whatever we might say to ourselves, the interpretation we choose is a matter of subjective preference. In some sense, it is even an ideological motivated impulse, not a rational and scientifically sound choice. Nothing is wrong with this per se, as science has always progressed by a trial-and-error approach. Hypotheses, conjectures, and sometimes even wild speculations can and did lead to positive results and groundbreaking discoveries. However, we must be aware of our (more or less unaware) assumptions and be able to discriminate between a model based on scientifically established facts versus one linked more to a speculation we like to advance for personal reasons. This is an ability and skill that nowadays seems to be increasingly rare.

One of the main motivations for a healthy skeptical attitude towards interpretations of QM is the fact that, as their denominations already spell out, they are mostly interpretations, not full-fledged theories that predict something new. The difference between an 'interpretation' and a 'theory' is that the former is descriptive of only something which is already known, with some different model or formal approach, but does not allow us to distinguish it experimentally from others by pointing at some new phenomena. There are a very few exceptions in which an interpretation also makes some prediction (such as the GRW objective collapse) but these are either so far disconfirmed or extremely difficult to test with present technology. Quantum interpretations are sometimes models of completely different ontologies and yet are empirically almost indistinguishable from orthodox QM and furnish no element that could either falsify or confirm it. The above-mentioned models or formal approaches to QM do not add substantial new insights or formulate new predictions. They tell us what we already know.

This is at odds with the history of science, which shows us that new ideas, concepts, or models ignite paradigm shifts only if these lead to new insights that the old theories or models were unable to predict. For example, the shift from the geocentric to the heliocentric model of the Universe combined with Newton's gravity is able to not only correctly and more easily describe the already-known planetary orbits (we will see next that this alone isn't a real issue in the geocentric context) but also predict completely new phenomena, such as the stellar parallax–that is, the slight shift in position, a strange seasonal wobbling of the 'fixed' stars in the sky. Einstein's theory of relativity was not just another interpretation of classical mechanics. Special and general relativity are reformulations that extend CM and also predict new phenomena, such as time dilation and space-time curvature, or explain anomalies, such as the perihelion shift of Mercury and other deviations from

CM, and which otherwise would remain totally unexplained in the frame of Newtonian physics. Relativity's new predictions were soon to be confirmed (for example, light bending or the slowing down of the muon's decay due to time dilation). And QP is also not just a re-interpretation of CP or relativity. As we have elucidated in the first volume, it came into being because Planck solved the anomaly of the ultraviolet catastrophe, which could not be explained within a classical context. Moreover, within a few decades, QM went way beyond that, introducing ideas, concepts, and a completely different formal approach that turned out to be one of the most prolific and successful theories in the history of science, leading to enormous theoretical developments and, ultimately, to the standard model of particle physics, not to mention the practical applications derived from solid-state physics.

Whereas, what, for instance, do the MWI or BM tell us beyond something which is already known? Besides conjecturing 'parallel worlds' which nobody knows how to verify, or almost metaphysical 'pilot waves' about which one wonders whether they will ever be observed directly as a separate entity because they are themselves responsible for everything we observe? Do these interpretations at least predict some new phenomenon, such as an effect for which conventional QP can't account, or some new particles? In fact, there have been attempts in this direction but, overall, after several decades of research, their potential to lead to some revolution and even some novelty in physics looks unlikely.

The reason we should take these interpretations with a 'grain of salt' is that the history of science has provided several instances of how models that made correct descriptions of what we know, but without making new predictions of what we don't know, always turned out to be wrong or, in the best case, if they couldn't be falsified, ended up as fossils unable to evolve (for example, Goethe's color theory of light).

One striking example we analyzed in this regard is the Bohr-Sommerfeld atomic model. With the electron circular orbits atomic model of Bohr, it is, indeed, possible to correctly calculate the hydrogen atom energy levels. With Sommerfeld's extension to elliptic orbits, even more spectral lines can be explained, at least at first approximation. It could, therefore, describe something already observed but could not predict any new atomic phenomenon. Quite the contrary; soon it became clear that this could not be the correct representation of reality, as too many unaccountable discrepancies and anomalies were observed in other atomic spectra. Nowadays, no physicist works with such a model. It has definitely been relegated to the status of historical curiosity.

Another interesting historical case was the '*caloric theory of heat.*' It was a widely accepted theory in the 18th century, positing the existence of a hypothetical invisible fluid, the 'caloric,' that could be neither created nor destroyed and that could be transferred from one body to another, responsible for the heat phenomena and process in matter, such as thermal expansion. It was quite a successful theory because it could explain several aspects of heat phenomena. Sadi Carnot (1796-1832), the French engineer and physicist famous for his theory of heat engines, could correctly derive its efficiency via his homonymous '*Carnot cycle*' by assuming the existence of the caloric. Also, Laplace used the caloric theory to estimate the speed of sound. He obtained a theoretical speed of 345 m/s against the experimental 337 m/s. [15] This is remarkably good agreement! And yet, nowadays, we know that the caloric theory is wrong. Heat does not have the properties of such an imaginary fluid. For example, with simple experiments, Joule could show that heat is not conserved in all thermal processes. More precisely, heat is a form of EM energy and, as any form of energy, it can be transformed into other forms of energy (for example, into mechanical energy, just as, the other way around, mechanical energy can be transformed into heat). Despite its partial success and its endorsement by great scientists, the caloric theory has been superseded by concepts and principles of modern thermodynamics.

However, my preferred historical example for illustrating how a model without any predictive power, but with good descriptive power, can nevertheless be wrong is the good old geocentric planetary model of Ptolemy. What historians and philosophers of science have frequently pointed out (but what physicists tend to forget) is that it is perfectly possible to accommodate, to any degree of precision, the apparent planetary motion of the Moon, the planets and the Sun on the celestial sphere with the epicycle theory of geocentric astronomy. Ptolemaic astronomers imagined a geocentric model with planets orbiting in a small circle with a non-uniform motion, the '*epicycle*', which in turn moves along a larger circle, the '*deferent*'. In this model the Earth was slightly off-center from the center of the deferent (that is, the Earth was almost at the center of the Universe). Thus, it was possible to describe, at least to a first order precision, the path of the planets, some of which, like Mars, displayed also that annoying retrograde movement from time to time. The question of why planets follow epicycles and what kind of 'force' or whatever kind of physical influence acts on them to 'guide' them along these paths remained unanswered.

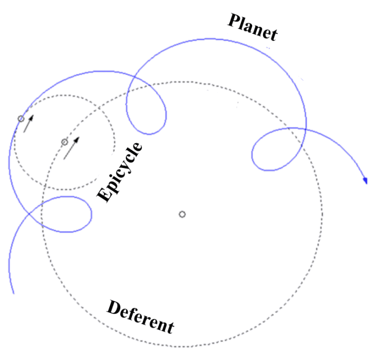


Fig. 31 The geocentric epicycles.

We won't go into the details of this theory, which, if analyzed in detail, can turn out to be quite complicated. The point of interest for us is that, with such a model, virtually every kind of orbital path in the sky, however complicated, can be described with any level of accuracy. In fact, by centering a second deferent moving on the circle of the first one (and eventually a third to the second, a fourth to the third, and so on), the planet moving on the epicycle with an appropriate radius and angular velocity will nicely trace any desired trajectory—not only all those curves observed in astronomy but also any function of interest, as, for example, an approximated square function. Fig. 32 shows how, with the addition of four epicycles with different radii and moving at different angular speeds, a square function can be approximated—or, also, the craziest ones imaginable, such as a point that traces a curve which writes "hello world," as in Fig. 33. [16]

This should, however, not come as a surprise to anyone with a background in first-year undergraduate math in physics or engineering. Any path constructed with a sufficient (eventually infinite) number of epicycles is nothing more than a function resulting from a *Fourier series*.

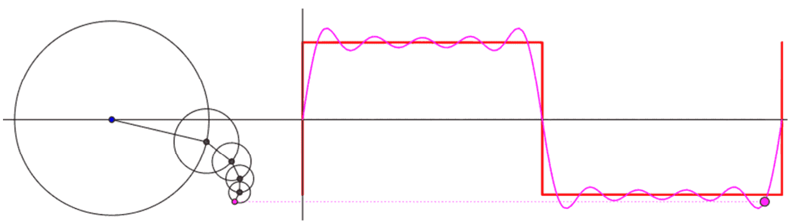


Fig. 32 Fourier series with four terms approximating a square function. ©Steve Phelps [17]

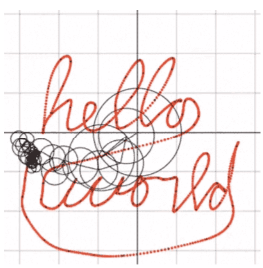


Fig. 33 Only phantasy limits epicycles!

For those who are not acquainted with the Fourier series, it may be said only that the French mathematician Joseph Fourier (1768-1830) was able to prove and construct a rigorous and elegant theory which shows how by adding a series of oscillating cosine and sine functions, each with an appropriate amplitude and angular speed, any periodic function can be represented. Or, to put it bluntly, any series of waves that have the appropriate frequency and amplitude and that overlap in space can interfere, giving rise to whatever kind of pattern results. It is only a matter of finding the appropriate coefficients to determine each amplitude and wavelength and, voilà, you get whatever you wish. Of course, these functions have nothing to do with planetary orbits but, because a circular motion can be decomposed into two oscillating modes of a sine and cosine function, it becomes obvious why the planetary motion of planets moving around the Sun can still be explained away by a geocentric epicycle theory at any desired degree of precision. In principle, there is no need to switch from the geocentric model to the heliocentric model.

The reason why nowadays celestial mechanics does not use any epicycle theory is, of course, because we know that the Earth moves around the Sun, but also because the heliocentric perspective makes calculations much easier. The 'geocentric interpretation' is only descriptive and not predictive of new phenomena. Ultimately, anomalies appeared that can't be explained if not in the context of heliocentrism. The anomaly in question is that of the above-mentioned stellar parallax, which could finally be observed only in 1838. In other words, almost three centuries had to pass after Copernicus's conjecture that the Sun is at the center of the solar system before science could finally prove with absolute certainty that his model was correct! Heliocentrism with Newtonian gravity not only described already-known facts or explained anomalous phenomena but also predicted new things. Most famously, it predicted the existence of a new planet that, until then, nobody suspected to exist—namely, Neptune.

I dwelled a bit on these pre-quantum historical anecdotes only to show how modern science tends to fall into the same trap. We must be careful in embracing any ontology or model that supposedly describes an objective reality if it describes the observational and experimental data alone. If a theory and its math, as was the case with the epicycle theory, Bohr-Sommerfeld atomic model or the caloric theory, correctly describes the phenomena and 'saves appearances,' this does not imply that we must believe it to be correct. A mere calculation tool that confirms our observations does not necessarily tell us something about an objective reality.

A partial analogy for the epicycle theory in QP can be drawn with Feynman diagrams. In Vol. I, we saw how interactions between quantum particles have nothing to do with the naive understanding we have at the macroscopic scale. For instance, a scattering process between two electrons has nothing in common with the scattering of two marbles. We must understand it as an interaction between two repulsive force fields. In *quantum field theories* (QFT), this process is described by taking another perspective: These forces are represented by Feynman diagrams which we like to interpret as the exchange of particles, either real or virtual. *Quantum-electro-dynamics* (QED) and *quantum-chromo-dynamics* (QCD), the modern quantum theories of particle physics that describe the EM interactions and the strong as weak nuclear forces, rely on a calculation procedure, the '*perturbative approach*' to Feynman's path integrals. Every interaction—that is, a scattering process—is not the result of the exchange of a single real particle but, rather, is expressed as the sum of a series of Feynman diagrams, each representing a virtual particle exchange. It is a similar principle of the Fourier series. Here, instead of epicycles, the perturbative terms are Feynman diagrams—that is, particle exchanges are summed up to approximate the real process we observe in the laboratory It should be considered only a partially applicable analogy because, admittedly, perturbation theory with Feynman diagrams could also predict new phenomena. It works, it perfectly matches the data, and it has become a powerful calculation tool for one of the most successful theories in the history of science!

However, because the epicycle theory had nothing to do with reality, suggesting even the exact opposite of truth, here we should be careful as well. History tells us that the fact that we found a calculation tool that nicely describes reality with a perturbation series must not lure us into the belief that each of the expansion terms has anything to do with the description of something real. From the ontological point of view, there are no such things as ‘particle exchanges’ and even less ‘virtual particles.’ These are ‘real’ just as the epicycles in the geocentric model of the Universe are ‘real.’

But in what sense does this relate to the interpretations of QM? Take, for example, the MWI. Here, one does not multiply epicycles or graphs but, rather, entire ‘worlds.’ The difference is that it does not sum them up but ‘splits’ them. Does this really make a difference when it comes to the question of whether it describes anything real? And frankly, this idea that the author, with all that he is and feels to be, is splitting in gazillions of parallel worlds every second, is a conjecture that, psychologically speaking, doesn’t sound very realistic, to put it mildly.

And what about BM? Bohm and de Broglie went the other way around; they did not multiply entities but instead posited the existence of a ‘guiding’ or ‘pilot wave’ and a ‘quantum potential’ that guides the particles along the desired path, simply to ‘save appearances’ again. It is the same as if the ancient astronomers had introduced some complicated force fields or gravitational potential that guided and ‘piloted’ the planets around the epicycles which, in turn, travelled along the deferent. Of course, that would have worked but it would have had nothing to do with reality. This is not only because we know that the geocentric system is false but also because such a force field does not exist and, if it did, would have nothing to do with Newtonian gravity anyway. And how much does the idea of conceiving of the wavefunction as having fluid-dynamical properties differ from the invisible fluid of the caloric theory?

The question is also why people are so desperately looking for interpretations of QM. QT works perfectly, so why bother? The problem is that the microscopic quantum properties and events do not meet our human macroscopic intuition, according to which the physical world appears as a Laplacian deterministic Universe where laws of local realism hold and everything is reducible to interacting 'marbles' that we call 'particles' and from which, at least in principle, everything could be derived, constructed, and explained. We refuse to admit that our human mental perception of the world in terms of deterministic and, at least in principle, always predictable phenomenon and well-defined properties of physical objects does not apply to the microscopic realm (and which, admit it or not, we more or less unconsciously continue to frame with a play of tiny 'billiard balls'). We therefore insist on trying to apply a classical Newtonian understanding of reality in the quantum domain because of an unaware anthropic reasoning. It is the same with superdeterminism, which is one of the typical loophole arguments that looks to be the weirdest, but still conceivably most logical, possibility left for maintaining a deterministic model of reality at any cost. In the author's view, though, this is a sign of desperation due to a refusal to give up one's own anthropomorphic preconceptions.

But, after all, this is how science works. It must consider, and test as far as possible, all the alternative explanations before definitively embracing one model. It is no more and no less than what the followers of Ptolemy did in trying to maintain the geocentric model because this is what our senses suggest upon looking at the rising and falling Sun. When confronted with this objection, several physicists reply that the deterministic and reductionist approach worked and therefore they do not see why it should be given up. We should not forget that, a theory can work in its domain, as Newtonian physics eloquently showed. It does not need QM or relativity for many purposes and situations but breaks down beyond a certain limit. Humans were sent to the Moon using classical celestial mechanics. So what? Similarly, determinism, local realism, and reductionism make sense on a certain spatial, temporal and energy scale and in some specific domain but break down in another one. The argument that a theory or an approach to reality 'works' tells us nothing about reality itself.

The author is not against new interpretations and theories that seek to make sense of QM. However, first of all, the lesson that history tells us is that any breakthrough in this direction is unlikely to occur if we are not willing to give up our human preferences about what Nature is supposed to be, based on an analytic mind that admits only the classical causal understanding stemming from the limited sensorial apparatus and conception of Homo sapiens. Secondly, as long as an interpretation of QM remains only that, namely, an interpretation that does not lend itself to falsifiability with experiments and that cannot make new predictions pointing to new phenomena, it will remain a highly doubtful endeavor. If it does not propose an empirical test and an experimental demonstration that distinguish it from other interpretations and theories, it does not serve as anything other than an intellectual exercise. If such things as parallel worlds, pilot waves, quantum potentials, or any sort of invisible ‘quantum fluids’ are mere abstractions and calculation tools useful only for saving appearances, but nobody can prove to exist having an objective reality, such as one can do with gravity or the EM and nuclear forces, then de Broglie’s and Bohm’s attempt, just as all the plethora of interpretations, will remain nice toy models which, however, have nothing to do with reality. These resemble much more the desperate attempt of an anthropocentric mindset that clings to an Aristotelian worldview multiplying entities such as epicycles and deferents instead of accepting the Copernican insight that the world is not as it appears to our senses and minds.

III. The standard model and beyond

1. Quantum Field Theory

We now have a sufficient background with which to discuss the '*standard model of particle physics'* (SM) and the lines of research that seek to build a unified theory of fundamental forces and that scientists are currently investigating. It builds upon classical mechanics, statistical mechanics, and especially the theory of relativity and quantum mechanics.

The enhanced version of QM is '*quantum field theory*' (QFT). The bad news is that its description is much too complicated to be summarized in a brief semi-popular science description here. It is covered in books with hundreds of pages that are full of calculations. The good news, however, is that its conceptual foundations are not very different from those of classical QM, which allows us to highlight its basics.

QFT extends the quantum laws of classical QP, which is still a non-relativistic theory, to that of '*special relativity*' (SR). Most importantly, QFT rests entirely on the notion of a '*relativistic quantum field*.'

In non-relativistic QM, the wave-particle duality, together with all the experiments on quantum-ontology that we illustrated in the previous volume and in this volume, made it amply clear how misplaced the concept of a point-particle is. However, so far, one could still stick with the point-particle idea by making calculations without bothering much about philosophical subtleties. In relativistic QM, however, it turns out that even this is no longer allowed. If one insists on the image of the single point-particle moving throughout space and time, nothing really works and one has come to terms with contradictions with the theory of relativity. For example, trying to extend non-relativistic QM to relativistic QM leads to the violation of unitarity and causality, which means that there is a non-zero probability that a particle can propagate from a position A to a position B faster than light. Meanwhile, if one extends the concept of a discreet point-like particle of classical QM to that of a continuous field, everything behaves fine. In QFT, the fundamental entity is not a particle but a field. What we visualize as a particle is the excitation of a field. To each kind of particle one associates just a quantum field which obeys a relativistic wave equation, that is, an extension of the Schrödinger equation to SR. A photon, an electron, or a muon is described by its photon, electron, and muon field, respectively. The photon, electron, or muon is an elementary quantum excitation of the EM field, the electron field, or the muon field. These fields are represented as vibrating 'ripples' in some region of space-time. This ripple is thought of as an 'excitation' or a 'displacement' which varies in time. (As an analogy, in

1D you can think of something like small vibrating strings while, in 2D, you can think of a vibrating membrane and extend this to a 3D space.) The simplest version of a quantum field is the '*scalar field.*' This is a mathematical scalar function which defines, for every space coordinate $x = (x_1, x_2, x_3)$ and for every instant in time t, a scalar function $\phi(x,t)$, which is a number in space at time t. To provide an example with which we are familiar, recall the standing modes in the black body cavity discussed in Vol. I. It can be thought of as a 1D transverse vibration of a string, which can be described by a scalar field $\phi(x,t)$ that measures the displacement of each point from equilibrium, at time t, of a small element of string around a point x. Similarly, Fig. 34 left provides an example of a 2D scalar field for some instant in time. (Grayscale/colours represent the magnitude of the field.) Mathematically, there is no limit to extending the quantum field to any number of dimensions.

Scalar fields represent only particles with spin zero. Because most particles are bosons or fermions which have a spin, something more is needed. The generalization to a field for particles with spin is, in principle, not very difficult. We know how spin in QM is represented by objects in vector notation, like the spinors.

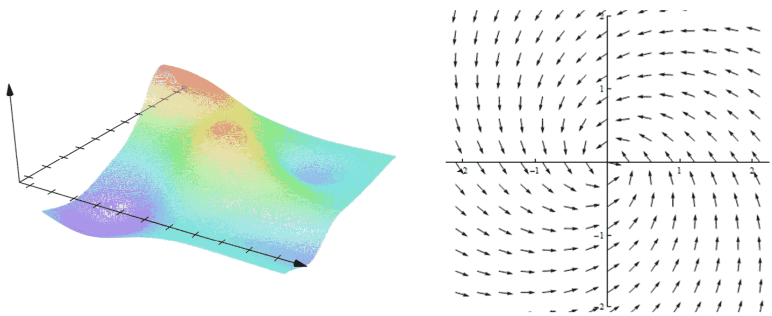


Fig. 34 Simple illustration of a 2D scalar field (left) and vector field (right).

Therefore, one extends the scalar field which identifies each point in space-time with a number to a '*vector field*' that, instead, identifies each space-time point with a vector (see Fig. 34 right). This is also why the *'Higgs boson'*, the only known scalar field and spin-less particle, and the photon which has spin, are also called a '*scalar boson*' and a '*vector boson*,' respectively.

There is a fundamental difference in how classical QM and QFT represent particles. In the former case, one deals with one quantum operator defined only on a point-like particle—that is, on one quantum oscillator in one point of space and one instant in time (a system with one '*degree of freedom*'). Whereas, in the latter case, the field represents an infinite number of oscillators on which an infinite number of quantum operators are also defined (a system with infinite degrees of freedom). Each oscillator is a 'quanta' with its quantum operator and which represents the particle of the classical quantum theory. This means that a single field can also describe many particles. This formal procedure of QFT is called '*second quantization*' or '*canonical quantization*' and became the conceptual and mathematical bedrock of the standard model (SM) of particle physics.

To understand what the SM is about, it is necessary to quickly recall the fundamental forces of Nature. We know about four fundamental physical forces. Possibly, there are more than four but these are what have been discovered so far. Two forces, we know about from everyday experience. A gravitational force is responsible for objects on Earth falling and having weight, for the fact that the stars and planets attract each other, and for forcing them into orbits or some more or less complicated trajectories, building up star systems, globular clusters, galaxies, and galactic clusters at cosmic scales. Then, of course, there is the EM force, responsible for the attraction or repulsion of electric charges and whatever causes radiation in the form of EM waves, such as light. What the gravitational force, mediated by a hypothetical particle called the 'graviton,' and the EM force have in common is that they can act at the range of large-scale distances compared to the micro-cosmos of nuclear forces. We can observe the light coming from galaxies billions of light years away, while these galaxies themselves are structures bound together by gravitational forces.

It is not so for the strong nuclear forces which act only at very small scales, i.e., much smaller than the size of an atom, at nuclear or femto-meter scales, which are about 10^{-15}m lengths—that is, of the order of a hundred thousand times smaller than the hydrogen atom. The weak nuclear forces act at orders of magnitude smaller.

This is why nuclear forces were discovered only recently in human history. They act on spatial scales that are much too small for human senses to perceive them. Nuclear forces are nevertheless extremely important to the existence of the Universe, at least the kind of Universe that we know. As we already discussed when covering the stability of matter, it is that force which holds the nuclei together. While the *'electroweak nuclear force'* is responsible for nuclear reactions and the process of atomic decay.

The SM of particle physics de facto describes the processes related to three of the four fundamental forces, the EM, the weak and strong force, in a single theoretical frame. This brought science to the verge of unification, as only the force of gravity is left out of the picture. The idea of bringing the four fundamental physical forces under the 'same umbrella' inevitably requires a unification of GR and QM, as two of these forces—gravity and EM—act in a general relativistic macroscopic setting, while the other two express themselves in a quantum micro-cosmos.

However, the goal of unifying GR with QM soon showed itself to be a formidable task. Therefore, physicists had to restrict their theory first to SR applied to classical QM. This means that one generalizes QM to that part of relativity which still does not contain gravity but nevertheless enlarges the quantum description of the particle world going beyond Newtonian dynamics. This gave birth to modern particle physics in the form of the SM.

For example, the Schrödinger equation, the equation we addressed in the context of atomic physics, still relies on a formulation without relativity. It was obtained by Schrödinger considering the expressions for the energy and momentum operators in a purely classical non-relativistic frame. Later, however, the same equation could be generalized to SR, first in 1926 by Oscar Klein (1894-1977) and Walter Gordon (1893-1939), and then in a more complete and developed form a couple of years later by Paul Dirac. The former equation, the so-called '*Klein-Gordon equation*,' describes particles relativistically but without spin, while the latter, called the 'Dirac equation,' includes particles with spin.

Dirac's equation not only did that, but it also predicted the existence of anti-matter—those kinds of particles we encountered in the pair-production and annihilation process. The Dirac equation was also an interesting historical case that outlines how difficult it sometimes is to make sense of the mathematical formalism. His equation, in fact, predicted the existence of particles with negative energy. It was not easy for him to make sense of it. What is an object that has negative energy values? He first interpreted this by conjecturing that the vacuum is completely filled with particles like electrons and that those which have a positive energy are the normal ones, while the negative-energy particles correspond to empty 'holes' in a sea of particles, the '*Dirac sea*'. Nowadays, we know that the interpretation in terms of anti-matter—that is, a form of matter which comes up with exactly the same particles as our ordinary matter but with opposite electric charge and spins—is the correct way to interpret his equation. This is because anti-matter has been produced and observed in particle accelerators.

Therefore, to sum up, there have been at least three levels of generalization.

First, from particles described by a classical quantum picture, to an extension of the theory that includes SR, first with a spin-less special relativistic expression and then including particles with spin, and whereby the existence of anti-matter became an almost natural prediction. The prediction of the theory, and the later empirical proof that anti-particles exist, showed that unifying at least SR with QM is possible. This, in turn, signifies that physicists were heading in the right direction (and, again, that a scientific theory is such only if it predicts something that can be tested or disproved).

The second level of generalization was that of switching from a discrete particle notion to a continuous field description. In this description, one no longer visualizes simply particles but, rather, assigns to each of them a field in space-time A scalar field is a field which tells us something about the probability of finding spin-less particles interacting somewhere in space-time, while a vector field does the same for particles with spin.

Fig. 35Sheldon Lee Glashow, Abdus Salam, and Steven Weinberg.

Finally, the third step was that to apply the second quantization—that is, applying an operator at each space-time point and switching from a description with one degree of freedom to that with infinite degrees of freedom. This formal approach together with SR gave birth to what is called '*quantum electro dynamics*' (QED). It describes the EM forces and weak nuclear forces unifying them in an '*electroweak* force' (or electroweak interaction - EW). In the 1960s, Sheldon Lee Glashow, Abdus Salam, and Steven Weinberg independently discovered that they could construct a theory of the weak force that included the electromagnetic force and received for this intellectual achievement the Nobel prize in 1979.

Force	Gravity	EM	Weak	Strong
Force carriers	graviton	photon	$W^{\pm}$ W^{-} Z^{0}	8 gluons
Particles affected	all known	electrically charged	all known	quarks gluons
Relative strengths	10^{-38}	10^{-2}	10^{-6}	1
Range	∞	∞	$10^{-18}m$	$10^{-15}m$
Theory	GR	← QED →		
			← QCD →	
		← SM →		
	← Quantum gravity? →			

Table 1 The four fundamental forces of Nature

At this point, physicists had some basic ingredients with which they could work towards an extended model of the particle world. This line of research was pursued for about half a century, between the 1930s to the end of the 1970s. To some degree, in some details, it continues today, and has given birth to what is known as the SM of particle physics. The SM can be described as one of the most successful theories of physics, if not of the history of science. Its predictions about matter and radiation were tested experimentally to the highest degree of precision and it is today considered one of the main pillars of modern physics. Let us see in the next chapter how it is structured

2. The Standard Model of Particle Physics

To get an idea of what the SM is about, we must first understand the overall representation and hierarchy of matter that emerged from it. As is well-known, matter is made of atoms and molecules. However, this is still a too-large-scale domain. The SM describes matter in its fundamental constituents from atoms downward, in terms of subatomic particles that are considered 'elementary'. Whether or not these particles are truly elementary is still unclear, but we will nevertheless label as such all the smallest to-date known particles.

The classification of matter according to the SM of particle physics is first of all divided into two main categories that you know well by now: fermions and bosons. Fermions are all those half-integer spin particles which make up matter, while bosons are integer spin particles, most of which are responsible for the mediation of forces among fermions. If you like, you can intuitively think of fermions as matter particles and bosons as force particles (with the exception of the Higgs boson).

Fermions are, in turn, subdivided into two other important categories: '*quarks*' and '*leptons*'. The former are those particles which make up the proton and neutron in the atomic nucleus. There are six types, or '*flavors*', of quarks (again, pysicists like to resort to bizarre analogies) and each possesses a fractional positive or negative elementary electron charge e.

The quarks' names – 'up', 'down', 'strange', 'charm', 'top', and 'bottom' – are obviously only labels that physicists like to assign to invisible objects. They do not mean anything that relates to visible colours. Protons are made of two up quarks and one down quark, while neutrons are made of two down quarks and one up quark. There are also other possible combinations which form several other particles, like the π^0 and π^+ mesons, which, however, are all unstable particles and decay in extremely short times.

Because of its minuteness, the mass of an elementary particle is, in particle physics, no longer expressed in kg but in $\frac{eV}{c^2}$. An eV ('*electronvolt*') is the energy that an electron acquires if subjected to an electric potential of 1 Volt. From Einstein's mass-energy equivalence $E = mc^2$ (see pair production in Vol. I), it follows that mass can be expressed as $m = \frac{E}{c^2}$, that is, proportional to an energy unit. Therefore, one can define a mass unit of

$1 \frac{eV}{c^2} = 1.783 \times 10^{-36} kg$, which is really small enough! An electron has a mass of 511 $\frac{MeV}{c^2}$ (the suffix M stands for '*mega*', that is, 10^6) or the proton has a mass of 938 $\frac{GeV}{c^2}$ (the suffix G stands for '*giga*', that is, 10^9) and is therefore about 1836 times more massive than the electron.

Please avoid the frequent misconception of mass being a measure of the 'hardness', 'impenetrability', or 'materiality' of an object. In physics, mass is a measure of inertia–that is, of how difficult it is to accelerate a body. As discussed at length in Vol. I, what causes an object to acquire its macroscopic property of solidity is not its mass but, rather, the microscopic interaction via one or more fundamental forces with other particles. See further the cases of neutrinos, dark matter, and dark energy.

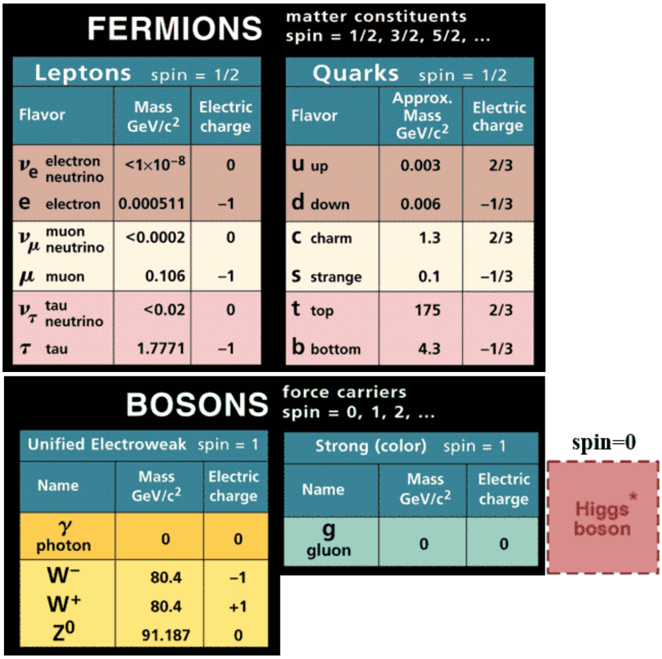


Fig. 36 Fermions and bosons according to the standard model of particle physics.

More generally, all particles made up of quarks are called '*hadrons*'. There are two families of hadrons: baryons (made of three quarks) and mesons (made of one quark and one anti-quark). Therefore, protons and neutrons are baryons, the only two stable hadrons, which build up the visible material universe together with the electron.

However, at least so far, no internal structure has been seen inside an electron, which is therefore considered to be elementary and composing part of the other family, that of ‘*leptons*’. Together with the electron, there are also five other leptons – namely, three neutrinos, a muon, and the τ particle. Of special interest are neutrinos, which are particles whose interaction with matter is extremely weak. They can traverse the whole Earth, and even entire stars such as our Sun, without being affected because they do not interact with other particles, nor by EM interactions or the strong force, but only with the EW force. This is the reason why these evanescent particles, with a very small mass, are also extremely difficult to detect.

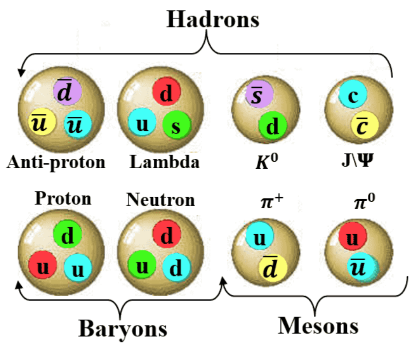


Fig. 37 Hadrons according to their quark structure.

Then we have bosons. Of course, you know that the typical example of the boson is the photon, which carries the EM force. Its interaction with matter in the frame of SR is described by QED. However, there are also those bosons which carry the strong nuclear forces, called 'gluons', named after the word 'glue' for the obvious reason that they glue quarks together to form the proton, neutron, and other hadrons. This is the domain of QCD. Gluons also have some type of 'charges', like the proton and the electron. However, they do not have electric charges; rather, they have strong force charges, prosaically called 'colours' by physicists. A peculiar feature of the strong force is that it acts in the opposite manner as the EM forces or the gravitational one. For the latter, the separation between charges or masses causes the EM and gravitational forces to decrease. Meanwhile, for the strong force, the attempt to separate quarks results in an increased 'gluonic' force that holds together the nucleons. An infinite force – that is, an infinite energy – is necessary to separate quarks. This is the reason why quarks have never been observed as separate particles and why their existence has been inferred only through indirect observation (the scattering process between nucleons). This effect is called ‘colour confinement’.

Then there are the $W^{\pm}$ and Z^0 bosons, which are responsible for the weak force and cause the nuclei or other leptons to decay.

Finally, the SM predicted the existence of the Higgs boson, which, after a long period of research that lasted about 40 years, was discovered in July 2012 at the LHC and for which Peter Higgs and Francois Englert were awarded the physics Nobel Prize.

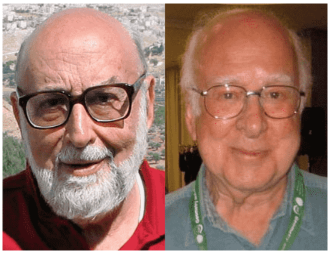


Fig. 38 Francois Englert and Peter Higgs.

The popular media like to call it the 'God particle'. The Higgs boson has, however, nothing transcendental or mystic about it. It is simply a label invented by a publisher for marketing purposes, to sell more of its books. The Higgs particle is a boson not because it carries a force but because it has spin zero, which is the overall definition of bosons. It also has no electric charge and no colour charge, and it decays in an ephemeral time lapse. Photons, the carriers of the EM force, are massless, just as all the carriers of the strong nuclear force, the gluons, are massless, whereas the '*Higgs mechanism*' explains why weak force bosons, the $W^{\pm}$ and Z^0, have mass.

Fig. 39 The 27 km circumference Large Hadron Collider (LHC) - Geneva (Switzerland).

A peculiar type of field is postulated – the '*Higgs field*'. It is believed that this Higgs field is non-zero in empty space and is a sort of background field that fills the whole Universe, like the zero-point vacuum energy. The theory predicts that this Higgs mechanism breaks the symmetry of the EW interaction and lets the EW bosons acquire a mass. The same field and mechanism also explain why all the particles we know have a mass. Initially, all particles are massless, but most of them interact with this background Higgs field, doing so by acquiring mass. The analogy frequently used to visualize this mechanism is that the Higgs field permeates the universe with countless Higgs bosons, acting like a fluid acts with friction on objects. A particle that suffers less friction while traveling through the Higgs field has less mass and vice-versa. What we call 'mass' is, therefore, a measure of the particle's interaction with the Higgs field and this, ultimately, is responsible for its inertia. Photons and gluons do not interact with it at all, and, therefore, have zero mass.

The Higgs boson was the last piece that was lacking in the SM of particle physics. Its discovery consecrated the SM as the best and most accurate theory of matter we have had until now. In fact, with this theoretical framework, the SM is able to unify three of the four fundamental forces of nature. The EM force, the strong force, and the weak forces are all seen as one fundamental interaction arising from the Higgs field. To put it in a simplified manner, we might say that the bosons of the weak and strong forces are the photons of the EM field which undergoes a *'spontaneous symmetry-breaking'* acquiring mass (more on this in the next chapter). In this sense, there is a unified vision, as these three forces emerge from the very same particles but behave differently according to their masses, if any.

This was, of course, only a very sketchy outline of the SM. However, it should give you an idea of what we know today about the basic stuff of which the visible Universe is made. However, despite being one of the greatest triumphs of theoretical and experimental physics of the 20th century, everything indicates that it is only the beginning of a larger intellectual adventure.

First of all, the most notably absent ingredient for an ultimate theory of physical reality is the force of gravity. The SM describes a Universe without gravity and despite decades and generations of physicists trying to come up with a theory that incorporates gravity, there seems to be something in the structure of the SM that stubbornly refuses such an extension. However, the fact is that, fortunately, we live in a world with gravity. This fact alone is sufficient to understand how the SM can be only a provisional theory, not the last rung of the ladder.

Furthermore, there are many other aspects and missing pieces in the SM that suggest its limitations. For example, nobody knows why the fundamental constants in our physical theories have just that value that they have. In the SM theory, there are something like 25 arbitrary fundamental constants, like, for instance, the '*coupling constants*' (the parameters describing the strengths of forces), such as the *'fine structure constant'* (a number about 1/137 that describes the strengths of the EM interactions) and that of the EW and strong forces. They are just there, and nobody knows where they come from. Theoreticians simply fill the gap by measuring these parameters and using them in their calculations, though their origin remains unexplained. Even more mysterious is that the value of these constants seems to be 'fine-tuned' simply to create a Universe in which life can emerge. Should some (perhaps all) of these constants have a slightly different value, the Universe would end up in a dark and cold place filled only with particles in a chilling almost-absolute-zero void without stars and planets. This is the famous '*fine tuning problem'* that gave rise to so many debates and, not uncommonly, shifted to metaphysical and finalistic speculations.

Then, the question that remains open is: Where has anti-matter in our visible Universe gone? With particle accelerators, we can produce or annihilate lots of particles and anti-particles, but always and only both together. There is no known physical mechanism which allows for the creation of particles without anti-particles. Theorists conjecture that at the time of the Big Bang, another symmetry-breaking mechanism must have taken place that allowed for the creation of a Universe made only of matter without anti-matter; however, the SM of particle physics has not been able to discover such a mechanism. Matter is just there, and anti-matter has disappeared for a reason we still have to figure out.

Moreover, astronomers tell us that the Universe must be filled with some mysterious and, until today not-better-specified, '*dark matter*' and '*dark energy*'. We know of the existence of dark matter (more precisely, one should call it 'obscure matter' because its 'darkness' has nothing to do with a colour) only because of its gravitational effects on the star dynamics in galaxies and among clusters of galaxies themselves. The existence of a dark energy that fills the empty space of all the cosmos is suggested by the fact that astronomical observations indicate that not only is our Universe expanding but that this expansion is accelerating. Except through the gravitational force, dark matter and dark energy seem to not interact with ordinary matter. Some conjecture about the existence of *'weakly interacting massive particles'* (WIMPs) that make up dark matter—massive particles which, however, are almost undetectable beyond their gravitational observational signature.

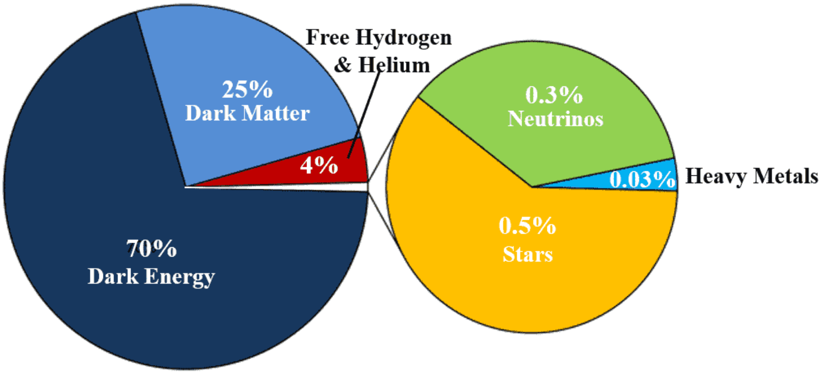


Fig. 40 The relative amount and types of matter in the Universe.

The relative amount of dark energy, dark matter, and visible matter makes the problem clear: The contribution to the mass of the universe is taken up by the dark energy by more than two-thirds. Then 25% is dark matter, while only about 5% is matter predicted by the SM (and of which only a meager 0.5% is matter in the form of stars and planets).

This can only mean one thing: What we see is only a little scratch on the surface of a universe which has almost certainly a much more profound and richer structure and content than our limited senses can see, inclusive of our most sophisticated detection instruments, like telescopes, microscopes, or particle accelerators. All of this, put together, clearly tells us that the SM cannot be the whole story.

3. Towards a theory of quantum gravity?

The problem that mid-20th-century physicists had to tackle was the simple fact that we now have two theories which work perfectly and have been experimentally verified several times: the theory of relativity (special and general) and QM. Today, there is no doubt that both theories are correct, at least in their essential parts. They have been confirmed experimentally and are perfectly consistent mathematically. Still, Einstein's relativity and QM are completely different descriptions of reality and present themselves as two complementary understandings. Relativity is a perfectly deterministic theory and, when probabilities are used, it is so only because of our ignorance. In principle, everything can be explained by classical statistical reasoning based on hidden variables. Relativity is a deterministic theory in which point-like particles have definite properties which move on a smooth space-time manifold with perfectly defined trajectories and states, and every correlation and interaction is represented inside the precepts of local realism, that is, the speed of light is a universal and omnipresent limit, a sort of ‘dogma’ of Mother Nature. On the other hand, QM is intrinsically statistical and a non-local theory (eventually without hidden variables). It does not allow for precise and definite properties; particles behave as waves, have non-local correlations, and can even be in a superposition of states.

And yet, despite being seemingly mutually exclusive, both approaches to physical reality work, are perfectly consistent, and have been experimentally verified with high accuracy. It was assumed that sooner or later one would win out over the other, and physicists are still speculating about which of the two theories should be considered more fundamental.

Will GR prevail, confirming our more intuitive understanding of the world in which objects have well-defined properties with an 'element of reality' which is independent of the fact of whether or not we are observing it, as Einstein would have opted for? Or will QM prove itself to be more 'fundamental' (whatever that might mean) and finally tell us that everything is indeterminate, inherently statistical, and non-local and that reality is contextual?

To date, nobody knows the correct answer for sure, but it might well turn out that neither the former nor the latter is more fundamental. Nature seems to think otherwise. It does not care about our limited understanding which wants to divide up things according to an idea such that one thing or the other thing must be completely true or completely false. It wants us to accept the fact that both things are true at the same time.

General relativity	**Quantum mechanics / QFT**
Deterministic	Non deterministic
Only point-like particles	Wave-particle duality
Definite properties and states	Properties and states are inherently indefinite
Space and time are dimensions	Space is an operator, time a parameter
Local realism	Non-local realism
Probabilities → ignorance	Probabilities → inherent 'randomness'
Statistics → hidden variables	Statistics necessary → no hidden variables
Describes gravity	Describes EM and nuclear forces

Quantum gravity?

Fig. 41 Distinctive aspects that distinguish GR from QM.

In fact, instead of thinking that one or the other will be de-selected, there is a third possibility – that is, both are two complementary aspects of reality, like the wave-particle duality, or, to put it in more symbolic terms, like the yin-yang complementarity, the dualistic concept of ancient Chinese philosophy which conceived of seeming opposites as complementary and interdependent realities. We are nowadays so far along with the theoretical development and its empiric evidence that we can fairly say that there is no sign that Nature favors GR or QM, one theory over the other. There is no mistake; both are true.

Therefore, since the time of Einstein, physicists have questioned whether there might be a more general theory that encompasses both relativity as QM, and where both are only two low-energy limits of a more extended theory, a QG theory – more or less like GR is a theory that encompasses SR, or like SR contains classical Newtonian mechanics on macroscopic scales and for low speeds, or like QM can be shown to converge to classical mechanics at macroscopic scales. A theory of QG also has the aim of describing the four fundamental forces as a unique and single force that manifests under different modes and ways in the low-energy physics in which we live. Here, also, we find a deep divide between GR and QM. The former describes the force of gravity well but seems to be at odds with the other three forces, while the latter has led us towards a unification of the EM and nuclear forces in the SM but seems to have no intention to leave much space for gravity. There is no reason, however, to believe that there can't exist a theory in which gravity is ultimately also a quantum phenomenon.

In popular media, QG is also called the 'theory of everything'. This expression is incorrect and we will not use it here because it is only a popular and misleading term. Even if we one day find a theory that explains GR and QM in a unique and unified frame, this theory alone won't tell us everything, such as, for example, how a human brain works (though there are speculations that QM might play a role in cellular activity, as we will discuss later). Additionally, the human, social, and spiritual conditions will hardly be explained by any theory of QG; even less will it be able to predict it. It is no more and no less than general relativity and QM actually already do. And houses, streets, bridges, cars, trains, and rockets will continue to be built according to the rules of good old Newtonian physics.

So, what is the rest of the story? As we mentioned before, the real thing that theoretical high energy physicists are after is a theory that must contain the SM while also being able to describe all four fundamental forces as the expression of an underlying unique '*super-force*' and that describes GR and QM as two limiting cases. So far, nobody has been able to present such a theory, and we don't even have any experimental evidence backing one or another theoretical model. Nowadays, there are several candidate QG theories that promise to do that; however, at least to date, they are mere mathematical speculations.

The theory that has gotten most of the attention in the last four decades is '*string theory*' (ST). It combines the quantum field theories of the SM with a new, more general theory based on GR and QM and, until recently, was considered by most physicists as the best candidate for becoming the future theory of QG, though there have been several other attempts of this sort.

ST is based on the idea that particles are not particles at all but, rather, extremely tiny strings (or, as in a later version, membranes). We are speaking here of objects which are about 10^{-35}m small – an inconceivably small length, the '*Planck length*'. Together with the '*Planck mass*' and '*Planck time*', they represent the '*Planck units*' which were, obviously, first defined by Planck, with a dimensional analysis, but which can be obtained by trying to find out at which scales the zero-point energy virtual vacuum quantum fluctuations become strong enough to form a virtual micro-BH. As outlined in the chapter on the zero-point energy in Vol. I, we know that, according to QP, empty space is not really empty. QFT envisages, and employs mathematically with success, the quantum superposition of energy levels above the ground state of the vacuum. One imagines this (with a naive and not quite correct ontology, but nevertheless helpful) as the vacuum being permeated by a continuously quantum 'fluctuating' space. At a microscopic-scale space, or more precisely space-time, is an eternally changing *'quantum foam'*. The smaller the region one considers, the greater are the energy quantum fluctuations. For space volumes which are as small as the Planck length and/or for time intervals as short as the Planck time, the energy density fluctuations become so violent, that the matter density according to Einstein's mass-energy equivalence is sufficient to form black holes (BH). And, because our current physical space-time theories break down when we are dealing with a BH space-time curvature, we can fairly say that at these scales, the known physical laws of QM and GR can no longer hold. Only a future and still-to-be-discovered theory of QG may eventually reveal to us how physical phenomena must be described at an even smaller scale. But as long as we have only the current physics, the Planck scale is the limit in spatial and temporal smallness beyond which we know for sure physics in its current state can't be used. Though the most powerful particle accelerators are also far from being able to test such tiny structures, the Planck scale is an important speculative concept of modern QFT. It is in this ultra-microscopic realm in which ST is modeled.

In ST, all known types of particles are just the same little string which is determined by its vibration frequency, its size, and its tension. ST is a mathematical model of the dynamics of these strings and which relates bosons to fermions by a '*supersymmetry*' transformation. Supersymmetry, also abbreviated SUSY, called promises to unify all the known particles – that is, fermions and bosons – by describing them as the two appearances of the same particle, namely, a string.

Fig. 42 shows how strings can be open, having loose ends, or closed. They can interact by connecting their ends and forming a new string. In Fig. 43, you can see the difference between an interaction among particles as in the SM and strings.

In the former, you imagine the two particles coming together into an infinitely small region of space-time and forming a new particle. This is what is called a '*singularity*', a point-like region of space containing a mass, and therefore mathematically equivalent to an infinite density. This causes, in the calculation, the emergence of infinite valued properties, which is a very undesirable side effect.

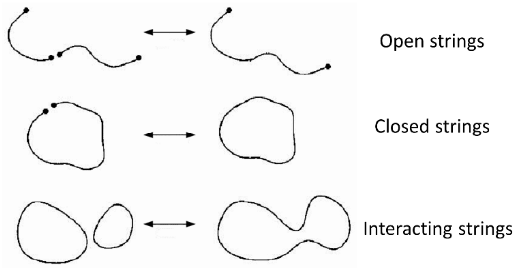


Fig. 42 Different types of strings and their interactions.

On the other hand, if you imagine strings instead of particles, the problem is solved in the first place: In the space-time region where the two particles interact, they form a new string with finite, non-zero dimension, and a whole bunch of calculations which were previously affected by annoying infinities now behave nicely.

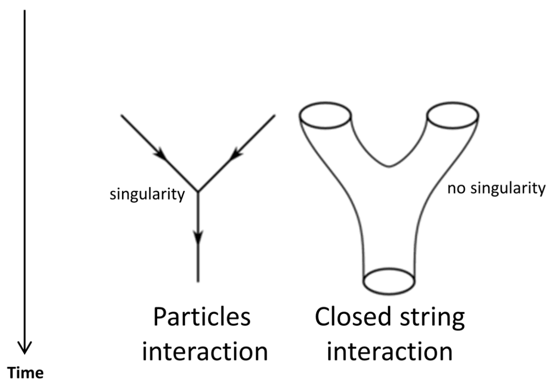


Fig. 43 Conventional vs. string interaction models.

Moreover, superstring theory no longer conceives of space-time as three space dimensions plus one time dimension but, rather, postulates the existence of another seven hidden dimensions at the Planck scale. We are aware of only four of the eleven total dimensions because the others are 'rolled up' or, as physicists say, they are '*compactified*' to extremely small scales that we can't notice. The string nature of particles, supersymmetry, and the extra dimensions make of ST an incredibly complex but rich theory that not only contains gravity as one of the fundamental forces (mediated by a boson with spin 2, the '*graviton*') but also requires it in order to be consistent.

Physicists and mathematicians began to develop string theory in the 1970s. In a couple of decades, five string theories were developed, which seemed to be unrelated to each other. However, later, in 1995, it turned out that they are the five aspects of the very same theory, which is now called the '*M-theory*' (from 'Mother-theory', or 'Matrix-theory').

These five theories are connected to each other through mathematical transformations called *'duality transformations'*, according to which they are 'dual' to each other in the sense that they are each a limiting case of the same underlying M-theory. The notion of duality is used in rather different contexts. Generally speaking, two physical theories are said to be dual if there exists a transformation, such that the first theory can be transformed into a second theory that looks just like the first one. A duality acts by interchanging the roles of two objects linking quantities that seemed to be separate (for instance, the electric with the magnetic fields, other strong and weak coupling strengths, large- and small-distance scales, etc.).

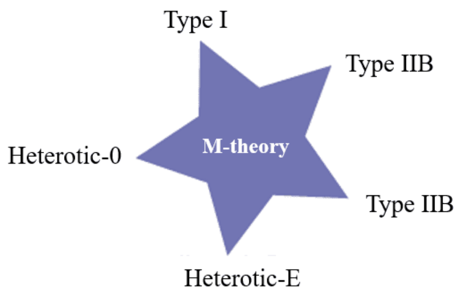


Fig. 44 M-Theory as the 'Mother theory' of all string theories.

This was only a brief sketch of a much more complicated theory that only a few physicists in the world know and can handle in detail. ST is not the only proposal for a QG theory. One could name several others, such as *'loop quantum gravity'*, '*twistor theories*', *'E_8 theory', 'geometrodynamics', 'asymptotically safe gravity', 'causal dynamical triangulations'*, and many other variations or attempts with strange and cryptic names – a multiplication of theories, conjectures, and hypotheses that share many of the similarities and shortcomings of the many interpretations of QM we discussed earlier.

Whether ST, or one of these alternative approaches, is indeed the correct description of nature remains a matter of debate. ST in particular has been criticized for many reasons. First of all, it deals with extremely small objects, which are not visible to even the largest particle accelerators, like the LHC. And they will remain directly unobservable for a long time because an accelerator as large as our solar system would be necessary to reveal them. The LHC reaches energies of about 14 TeV (1 TeV $= 10^{12}$eV – T stands for 'Tera'). Nature is capable, with some astrophysical monsters in deep space whose location and identity no one knows, of producing particles that reach energies up to 10^{22}eV – that is, about a hundred million times the LHC. The energy domain of strings is at the Planck energy scale, which is about another million times greater than that, at the order of 10^{28}eV. Therefore, ST describes a world which is totally out of reach and inaccessible to us. Because of this, many wonder if it can be considered science at all, as it seems to not be falsifiable. Fortunately, however (and this sets ST aside from being just another interpretation of QM), there are several low-energy versions of it that also make predictions which could be tested and are in the reach of the LHC.

However, string theories suffer from another severe drawback, namely, that they do not solve the problem of the parameter freedom of the SM. Quite the contrary; ST adds even more free parameters. There is nothing like a single ST, and there is an almost infinite number of ST versions. The lack of uniqueness of predictions due to the large number of solutions is such that one can accommodate every observation by simply *'fine tuning'* the internal parameters of the theory. This is also a highly doubtful approach to doing science criticized by many physicists. In principle, and this is the weakness that ST has in common with the long list of quantum interpretations, a *'string landscape'* of choices of parameters exists which allow the theory to predict whatever kind of Universe (something the author finds dangerously reminiscent of the epicycle approach!). That's why, nowadays, the '*Multiverse*' theory is so in fashion. We might live in a Universe which is only one of the many (perhaps infinite) possible versions inside a much larger Multiverse and where each of these Universes realizes only one particular set of physical constants. The problem with this, again, is that, first of all, the question arises: If a theory that is not a 'theory of everything' but, rather, is a theory that predicts everything we would like it to predict, is it a scientific theory at all? Science is based on precise predictions that can be verified with a 'yes' or 'no' answer, not on theories that tell us what we would like reality to be. Secondly, if other Universes really exist, these are completely disconnected from our Universe and nobody knows, not even in principle, how their existence could be tested empirically – yet another problem with falsifiability. The Multiverse hypothesis has, therefore, a sort of metaphysical taste that many physicists dislike and reject as unscientific.

Also, ST, similar to many interpretations of QM, makes the attempt to reintroduce a sort of quantized local realism. It posits at its foundations the existence of objects with definite properties, like a length and a tension, moving along well-defined trajectories, and then quantizes it, just as in QM one posits point-like particles and quantizes them. This goes against the suggestion of Nature (read: Heisenberg's uncertainty principle in its quantum random non-local realistic interpretation) to consider the quantum microphysical object not as individualized particles or strings or any other kind of geometrical objects with well-defined boundaries moving along precise paths but, rather, as non-separable, indeterminate, and fuzzy entities that have nothing to do with the naive models and pictures our mind suggests in the first place.

Will ST nevertheless succeed? Physicists are anxiously awaiting some signal from the LHC and other accelerators or astronomical observatories around the world, such as the discovery of a new SUSY particle predicted by ST or the detection of an anomaly that might indicate some 'new physics' that goes beyond the SM. However, so far, as of 2020, only the SM stands alone in all its glory, with a large desert surrounding it. Skepticism about ST is growing.

Actually, a debate is raging as to whether all these efforts put into ST, as also with other versions of QG, were effectively focusing on anything that has to do with physical reality at all. Recent theoretical insights also suggest that the model of an accelerated expansion of the universe caused by dark energy and ST are two mutually exclusive theories. The former or the latter must be correct but both cannot be true. And because the former already has some observational data that backs it, one is inclined to believe that ST must be an incorrect theory. In the last three or four centuries of science, there has never been a theory on which so many people worked for so long a time without tangible results. And slowly but steadily, a 'nightmare scenario' seems to have become reality: For the last four decades, thousands of physicists have devoted their entire careers to a highly speculative theory and a plethora of abstract mathematical constructs that might have nothing to do with real physics and the world we know. This also led to a discussion, at a more sociological and political level, of if and how some branches of modern science are still doing serious work and, eventually, what should be done to avoid other nightmare scenarios in the future. (See for example [18] or [19] and also the author did extensively write on this [20].)

IV. Quantum mechanics and information

1. Quantum teleportation

A mindboggling and interesting quantum effect is 'quantum teleportation'. This is something that was first conceived of by a group of scientists in 1993 and is reminiscent of the Star Trek movie, in which people and material objects are dematerialized and 'beamed' towards distant places almost instantly. It is an extremely comfortable type of teleportation system that we thought would forever remain a wild fantasy of sci-fi writers and filmmakers. And, in fact, we are still far from realising the teleportation of entire objects like a human body, or even a grain of sand. However, modern technology now allows us to realize it with single elementary particles like photons, or even with single atoms. This demonstrates that teleportation a la Star Trek might not be so unlikely after all and that it is, at least in principle, not forbidden by the known physical laws.

Considering the fact that teleportation has become a field of research in the context of quantum information, and also considering the kind of worldwide research going on nowadays in an attempt to build quantum computers, it might be a good idea to introduce here the basic concept of the 'qubit', opposed to the classical information unit with which we are all accustomed, which is the bit. In what sense does the qubit differ from the classical bit?

As everyone knows, classical bits—those with which our computers currently work—are simply cells of memory which can attain only two types of digital states: 1 or 0, 'on' or 'off', 'yes' or 'no'. Special emphasis must be set on the 'or', as they cannot be 1 and 0 at the same time.

Whereas, as we have amply discussed in the previous sections, in QM, quantum objects like elementary particles, photons, and electrons, as well as even larger objects like atoms, molecules, etc., can be in a superposition of states. If, for instance, we identify the two possible 1 and 0 states as the spin of an electron, we can agree by convention that a 1 bit-state corresponds to the spin up and a 0 bit-state to the spin down. This is just an example. Similarly, we might agree to define the 0 bit-state of an atom as its energy ground state, the 1 bit-state as the same atom on an excited energy level, and so on.

Let us keep general and speak about two quantum states and label it with the Dirac notation of $|1\rangle$ or $|0\rangle$. The decisive difference that we discovered in terms of quantum superposition is that a particle can be in both states $|1\rangle$ and $|0\rangle$ at the same time. Previously, we considered there to be only an equal 50% probability that one or the other state would manifest in a measurement.

We can, however, extend this writing to the state vector as:

$$|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle,$$

$$\text{with} \qquad |\alpha|^2 + |\beta|^2 = 1. \qquad \textit{Eq. 14}$$

where the modulus square of alpha and beta gives us the probability of finding the particle in one or the other state, respectively, and the sum of the modulus square of both provides unity, as the probability of finding it in either state is certain. It is nothing other than an extension of the previous equations. (Previously, we were restricted only to $\alpha = \beta = \frac{1}{\sqrt{2}}$.)

Eq. 14 represents the equation of a circle. This circle is represented graphically by an arrow, the state vector, that points to what is called the 'Bloch sphere', as in Fig. 45. The classical bit can place the state vector at only two points: the north or south pole of the Bloch sphere, the 0 or 1 state, while if we span all the α and β coefficients, as in the case of the qubit, the vector would span all the circle lines from the north to the south pole. We won't go into the mathematical details of this representation (which, by the way, implies another angular phase parameter that spans the entire spherical surface), but what is important to notice is that a qubit is really a very strange type of bit. It makes no sense to ask if the memory cell of the quantum computer is 1 or 0; it is in all the possible states in between, and only when you read it out will you get one or the other answer.

While the classical bit can be in only one or the other state, the qubit is in one AND all the other possible states determined by all the values that the α and β coefficients can take. This property gives tremendous computing power to quantum computers because they could, in principle, compute a much larger set of possibilities spanned by the α and β coefficients at once, while the classical computer must process each possibility one at a time.

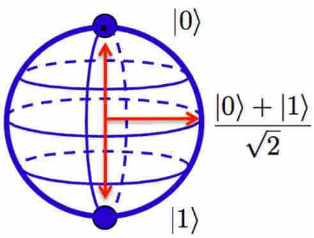


Fig. 45 The Bloch-sphere.

What is relevant to quantum teleportation is the entanglement of qubits. In fact, suppose that we label two particles as A and B. Then, if you followed what we have said so far about quantum superposition and the entanglement of fermions and bosons, as well as how the state vectors are represented with the Dirac notation, you should know how to read the so-called '*Bell-states*' and understand what they represent:

$$|\Psi^{\pm}\rangle_{AB} = \frac{1}{\sqrt{2}}(|0\rangle_A|1\rangle_B \pm |1\rangle_A|0\rangle_B), \quad \textit{Eq. 15}$$

$$|\Phi^{\pm}\rangle_{AB} = \frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_B \pm |1\rangle_A|1\rangle_B). \quad \textit{Eq. 16}$$

These tell us how the entanglement of these two particles, each having two equally probable possible states 1 and 0, must be represented in one of its possible entangled states.

In the first case, particle A and B are anti-correlated: If, upon measurement, particle A turns out to be in state 0, then the measurement of particle B must give 1, and vice-versa. (Recall how those with a positive sign could be, for example, the type-II entangled photons.)

Then, in the second case, particles A and B are correlated: If one gets particle A in state 0, then particle B must also turn out to be 0, and vice-versa. (Again, recall how those with a positive sign could be, for example, the type-I entangled photons.)

There is also the possibility of having the minus sign. As you recall from the lectures on the bosons, fermion, and Pauli exclusion principle, we explained what that means. On that occasion, we dealt with fermions which have no exchange symmetry; exchanging the order of two particles leads to a phase difference in the expression of the state vector, which is expressed by the minus sign.

The Bell-states express a set of four possible states of what is called a *"maximally entangled pair of qubits"*. They are called Bell-states — in honor, of course, of John Stewart Bell.

Let us now consider three particles, as shown in Fig. 46: Particles A and B, which are produced and entangled by what is called an '*EPR source*', in a similar way as it has been described in the section on the EPR paradox and the corresponding experiments therein, with the only difference being that we are not considering only anti-correlated particles but also correlated ones. Then let us also consider another third particle, C, which is not entangled with any other particle but which is in a superposition of states as:

$$|\Psi\rangle_C = \alpha|0\rangle_C + \beta|1\rangle_C \,. \quad \text{Eq. 17}$$

Therefore, we will intertwine the concept of superposition with entanglement here.

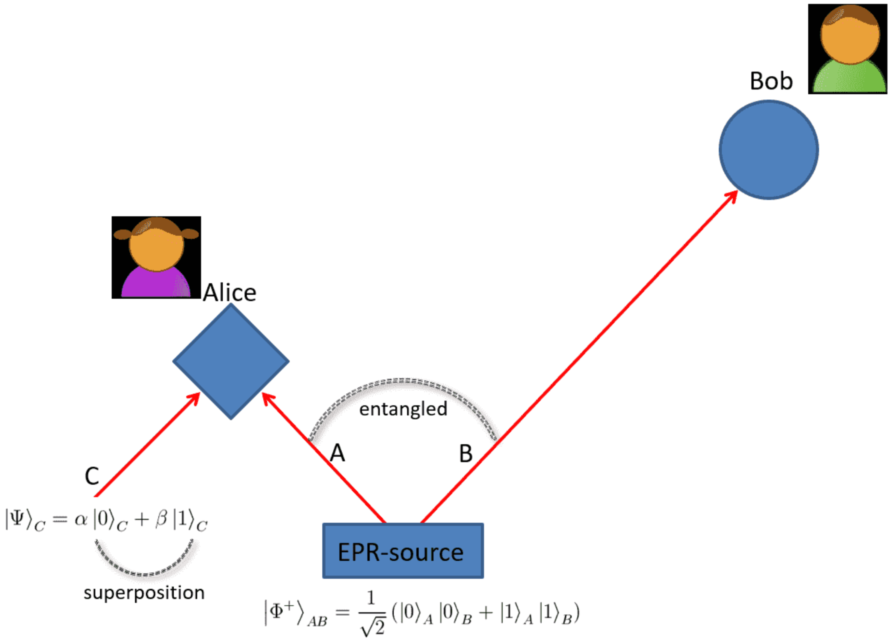


Fig. 46 Preparation of the quantum system before teleportation.

Now, consider Alice making measurements on particles A and C together, while particle B is sent at a very far distance to Bob. To be sure that there is no possible interaction between Alice and Bob, let us assume that they are separated from each other by light-years.

To formally see how teleportation is achieved, let us first make some formal considerations with a little algebra.

Consider one of the four Bell-states. Say that particles AB are entangled according to the Bell-state $|\Phi^+\rangle_{AB}$.

Then the entire system comprising the entangled pair AB and Alice's particle C, before Alice and Bob have measured anything, is in the product state given by:

$$|\Psi\rangle_{ABC} = |\Phi^+\rangle_{AB}\,|\Psi\rangle_C. \quad \text{Eq. 18}$$

Now it is time to take a pencil and a piece of paper and make some quite simple calculations (at times a bit clumsy, but it is standard algebra) in the following steps.

First, see how Eq. 18 can be expanded by explicitly inserting (the positive version of) Eq. 16 and Eq. 17, which leads to:

$$|\Psi\rangle_{ABC} = \frac{\alpha}{\sqrt{2}}(|0\rangle_A|0\rangle_B|0\rangle_C + |1\rangle_A|1\rangle_B|0\rangle_C) + \frac{\beta}{\sqrt{2}}(|0\rangle_A|0\rangle_B|1\rangle_C + |1\rangle_A|1\rangle_B|1\rangle_C). \quad \textit{Eq. 19}$$

This is the quantum state of the system of all three particles as a whole.

On the other hand, from the expressions of the Bell-states, we also obtain the expressions of what the entanglement between particles A and C would look like. Let us do that in preparation for the measurement that Alice will perform, by entangling particles A and C.

At this point, still from Eq. 16 and Eq. 17 (replacing the label B with C in the Bell-states), check that:

$$\frac{1}{\sqrt{2}}(|\Phi^+\rangle_{AC} + |\Phi^-\rangle_{AC}) = |0\rangle_A|0\rangle_C$$
$$\frac{1}{\sqrt{2}}(|\Psi^+\rangle_{AC} + |\Psi^-\rangle_{AC}) = |0\rangle_A|1\rangle_C$$
$$\frac{1}{\sqrt{2}}(|\Psi^+\rangle_{AC} - |\Psi^-\rangle_{AC}) = |1\rangle_A|0\rangle_C$$
$$\frac{1}{\sqrt{2}}(|\Phi^+\rangle_{AC} - |\Phi^-\rangle_{AC}) = |1\rangle_A|1\rangle_C \;.$$

Inserting these back into Eq. 19, one finally obtains:

$$\begin{aligned}|\Psi\rangle_{ABC} = &\frac{1}{2}|\Phi^+\rangle_{AC}(\alpha|0\rangle_B + \beta|1\rangle_B)\\ &+\frac{1}{2}|\Phi^-\rangle_{AC}(\alpha|0\rangle_B - \beta|1\rangle_B)\\ &+\frac{1}{2}|\Psi^+\rangle_{AC}(\beta|0\rangle_B + \alpha|1\rangle_B)\\ &+\frac{1}{2}|\Psi^-\rangle_{AC}(\beta|0\rangle_B - \alpha|1\rangle_B) \;. \quad \textit{Eq. 20}\end{aligned}$$

It looks complicated but the advantage of this expression is that we can clearly separate the particle's quantum state in the hands of Alice (the kets with subscript AC) and the quantum state of the particle that will reach Bob (the parenthesis of kets with subscript B). In more technical terms, one says that the system's state vector has been written in Alice's basis.

This is the three particles' state before Alice makes any measurement: the so-called '*Bell-state measurement*' (BSM). The BSM is a quantum mechanical measurement in which two qubits are entangled. If Alice performs a BSM on particles A and C, this entangles them. Before Alice's BSM, A and C were not entangled; only A and B were.

So, *before* Alice's BSM, the system as a whole was in this state, as written with the state vector of Eq. 20, as a sum of four terms.

After Alice has performed a BSM entangling qubits A and C, she reduces this state vector onto one of the four terms of Eq. 20 (all of which have a 1/2 coefficient in front of it, because there is an equal probability that the BSM will project it on one of the four possible states, that is, $\left|\frac{1}{2}\right|^2 = 0.25 = 25\%$ chance). Therefore, once Alice has measured the qubits A and C in the Bell-basis, the state of the whole system 'collapses' and the entanglement between particles A and B—that is, the entanglement between Alice and Bob—no longer exists, while the entanglement between particles A and C remains.

To see how quantum teleportation occurs, note that the quantum state of particle B in Bob's hands is now described by the alpha and beta coefficients, which previously described the quantum state of particle C in Alice's hands! That is, the quantum state of particle C near Alice has been teleported at the time of Alice's BSM, from particle C to Bob's particle B. And, as we know, this occurs instantly even throughout light-years of spatial separation.

On her side, Alice obtains one of the four states, each with 25 percent probability. Which one of the four states is a completely random process according to QM is not something that can be known in advance.

And still, a small operation remains to be done before we can be sure that we have exact teleportation. That is because, so far, the teleportation may still not be perfect, as Bob will also have an equal probability of obtaining one of the four states on particle B, which must not necessarily coincide with the case Alice got. To reconstruct exactly, with particle B, the state which was formerly that of particle C according to Eq. 20, he must also know which state Alice obtained.

This can be done simply by Alice's using a classical telecommunication channel, a telephone or a radio contact, through which she tells him which state she has measured.

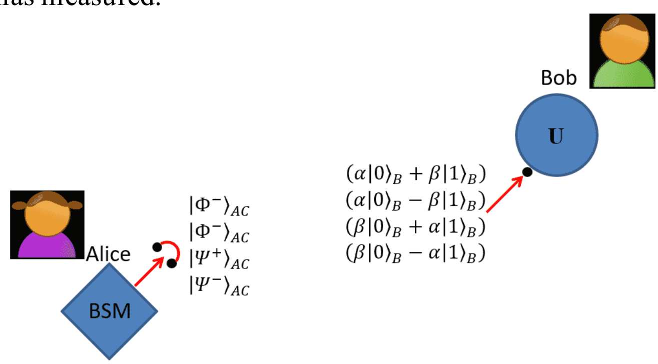


Fig. 47 The three particle ABC quantum system after teleportation.

Say she measured $|\Phi^+\rangle_{AC}$. Then Bob has to do nothing, as he already has the correct signature and order over the alpha and beta coefficients, and teleportation is complete.

On the other hand, if Alice obtains the Bell-state $|\Phi^-\rangle_{AC}$, then Bob knows that he must perform a little operation on his particle, namely, a *'unitary transformation'* which preserves the length of its state vector but changes its phase (we labelled that with the letter U in Bob's box) whereby he must change the sign on the beta coefficient (recall the definition of the unitary operator from Vol. I, chapter on the state vector and Schrödinger equation).

In the third case, if Alice obtains the Bell-state $|\Psi^+\rangle_{AC}$, Bob must perform a unitary transformation whereby the order of the two coefficients, alpha and beta, associated with the states 0 and 1 must be inverted.

Finally, if Alice obtains the Bell-state $|\Psi^-\rangle_{AC}$, Bob will have to invert the coefficients as in the previous operation but must assign a negative signature to alpha.

With this, the quantum teleportation process is complete.

Confused? To clarify some aspects and put things in their proper context, let us note two things.

First, to learn what he must do, Bob has to wait for Alice's call through a classical communication channel. Therefore, it is impossible to realize complete teleportation at a speed faster than light. Bob must still wait until Alice’s message arrives on a line which can, at best, transmit information at the speed of light but not faster, according to the theory of relativity.

Secondly, and most importantly, the question is: What has really been teleported? What happened is that the quantum state of a particle (in our case, particle C) has been copied exactly to the quantum state of another particle (to Bob’s particle B). It is simply an exact information scan, the information which makes up the physical object, that has been teleported, and not the physical object itself. Moreover, this implies the destruction of the quantum state of particle C. It is not a real ‘copy’ because the original quantum state of particle C has been destroyed, as particle C is now entangled with particle A. In fact, there exists a so-called ‘*no-cloning theorem*’ in quantum information technology that forbids the real copying of quantum states from one to another or many particles, maintaining intact the original state. There is no contradiction with the present laws of QM if we recall that this is a ‘copy and destroy’ operation at the same time. Only one particle is left with the original quantum state. So, it must be emphasized that, first, there is no real teleportation of matter but only of the quantum state of one particle to another. Secondly, the original particle's quantum state, which is copied light-years away, will have changed completely.

On the other hand, if you have read up to this point of the book, you should realize that, in QP, the distinction between a particle and the information about its properties and physical state is not entirely clear. To teleport the quantum state of a particle could be considered the teleportation of the particle itself. It is not so clear if we can really distinguish between a material particle and the quantum states that describes that particle. In fact, recall that when two particles are entangled, this means that both particles are with Alice AND Bob at the same time. The entangled duo of particles forms a unique and indistinguishable whole, before measurement. So, when two particles are entangled, where is the matter, the material aspect of the particles? According to QM, we must consider matter at the same time in both places, or even spanned all over space without distinction and real spatial separation, before measurement. At the root of all this is the principle of indistinguishability which is, as we outlined in the section on indistinguishability, much more profound and subtle than the classical conception of indistinguishability. Only at the instant when teleportation occurs (Alice's BSM) does the state vector collapse and we nicely have, again, two particles in two places: matter in two distinct places. So, it is not entirely clear if we are allowed to talk about the teleportation of mere information or if we should consider it, de facto, also a real teleportation of matter itself. The interpretation is up to you.

Of course, this is not mere speculation. Otherwise, we would not mention it here. This kind of quantum teleportation has been realized experimentally with photons. It has been shown how the teleportation of photons could be realized over a distance of 143 km in the Canary Islands. [21] A group of Austrian physicists, under the direction of (the already mentioned) Anton Zeilinger of the University of Vienna, used a laser beam attached to the telescope of an astronomical observatory on one island, which transmitted the photons and signals to another island. But, in principle, QM does allow any distance and any kind of particle quantum teleportation. The important point is that it has been shown that quantum teleportation is not just a sci-fi phantasy but, rather, an established experimental fact.

If we would extrapolate this to a quantum teleportation to larger bodies (say, a human body), this implies that an entire organism would be destroyed completely and rearranged and reconstructed particle by particle for each quantum state elsewhere. The original body, however, would probably have to die and dissolve. Personally, I would not like to undergo this kind of teleportation or 'beaming' process a la Star Trek. And this raises metaphysical questions. Is the teleported body still me? Is the soul, if it exists at all, teleported too? Would that 'I' that makes me feel 'me' be teleported as well?

Obviously, I have no answer and gladly leave these questions to you!

2. Quantum computing

A couple of decades ago, the aim to build *quantum computers* (QC) opened a new and exciting field of research. Quantum computing is based on the idea of using quantum mechanical principles to build a completely new type of computer.

With 'classical computers', one means just those computers that still every one of us uses daily and which are machines that operate on a set of stored strings in form of 'ones' and zeroes', commonly known as 'bits'. These work with registers. A digital register made up of two classical bits can store four digital numbers, namely (0,0), (0,1), (1,0), and (1,1), which we humans can simply label as 0, 1, 2, and 3. If the register is composed of three bits, it can attain eight states—(0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), and (1,1,1)—and could represent the numbers from zero to seven or any other set of symbols (such as letters, punctuations, operators, etc.). In general, a register with n bits can attain 2^n possible distinguishable states. By 'distinguishable', we mean that the register can store one of these states/symbols at a time but not all of them at the same time. Seems obvious, doesn't it?

It was Richard Feynman who, in 1982, first recognized that quantum mechanical phenomena can, in principle, be applied to the building of special kinds of computers capable of performing calculations that classical computers can't, namely, by considering that in QM, bits could also be in a superposition state.

We have already encountered the quantum bit, the qubit, in the chapter on quantum teleportation. It can be practically realized in many ways; for example, with photon polarization, the spins of particles or a nucleus, the two energy states of an atom such as a trapped ion or whatever kind of quantum system can attain only two states (up-down, V/H polarization, on-off, etc.). And we also know very well that, if properly prepared, a single quantum particle (a single qubit) can be in a superposition of states. This implies that a single qubit can be in state 0 and 1 at the same time or, more generally, a quantum register of n entangled particles in superposition forms an overall 'coherent' quantum state that can attain 2^n classical states at the same time.

This makes a quantum register an interesting device because, in contrast to the classical register, it can store all the possible states at once. Of course, once the register is read out, which means we perform a measurement of the qubits, the system 'de-coheres' due to state collapse and displays only one of the possible states, just like its classical counterparts. One might then question: What's the point of having a device that stores many states at the same time if, at the end of the line, we are allowed to read out only one? The trick is to induce interference between the qubits and get a result by coupling another qubit that doesn't take part in the calculation, with the function of the latter being to be read out by collapsing its quantum state but without interacting directly with the quantum register itself. To make this clear, let us proceed step by step.

Building blocks of classical computers are made of logic gates, and QC is no exception. Of course, *quantum logic gates* (QLG) differ from classical ones insofar as they can eventually have bit entries in superpositions and that can be entangled with other gates as well. In very general terms, a QLG is a '*black box*' or '*oracle*'—that is, a device whose internal structure we do not necessarily know, though do know how it acts on N input qubits giving an output according to some logical rules. An important aspect is that QLG must be unitary operators, which means reversible. Given a generic QLG labeled U, its reversibility is formally expressed by $|U|^2 = U{\cdot}U^* = I$ (just as we used to do with the modulus squared of the wavefunction) and which, loosely speaking, means that when applied twice, it outputs the input signal—that is, it works like the identity operator.

For example, one of the simplest gates is the '*negation operator'*, or '*NOT gate'*, which acts like shown in Fig. 49: It inverts the input signal.

Of course, applying it twice will send the quantum states of the qubits $|0\rangle$ and $|1\rangle$ into themselves.

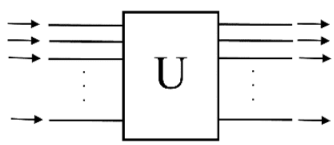


Fig. 48 Every QLG acts like a unitary operator on its input qubits.

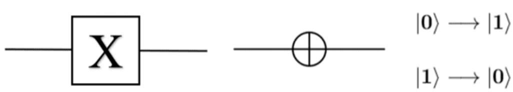


Fig. 49 The quantum NOT gate, its symbols and how it inverts the input.

One of the most genuine quantum gates, which has no classical analogue, is the Hadamard gate. It maps the single qubit basis state into the superposition state vector, as shown in Fig. 50. The *'Hadamard gate'* is responsible for the most fundamental quantum computation: It sets the eigenstates of the qubit into quantum superposition, as shown in Fig. 50.

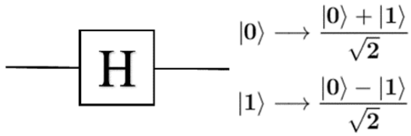


Fig. 50 The quantum Hadamard gate.

The Hadamard gate must act like a unitary operator as well when applied twice and map the basis states onto themselves—that is, act like an identity operator. (This is formally shown with matrix manipulation, but the reader who likes to shuffle a bit with algebra can verify this simply by applying H twice to the above state vectors.) Or, formally:

$$HH^*|0\rangle \rightarrow |0\rangle \text{ and } HH^*|1\rangle \rightarrow |1\rangle.$$

The same fits with the '*phase shift gate'*. It leaves the basis states unchanged but shifts the phase of $|1\rangle$ (relative to the input) by an angle ϕ, as shown in Fig. 51.

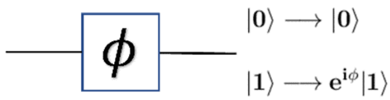


Fig. 51 The quantum phase shift gate.

Graphically, this can be visualized by moving the state vector on the horizontal circle of the surface of the Bloch-sphere we introduced in the quantum teleportation chapter (see Fig. 45).

Other single-qubit QLGs exist but the above suffice for our current purposes. Therefore, let us make some examples of two-qubit QLG.

A variation of the NOT gate is the '*controlled NOT'*, or CNOT gate, which differs from the previous one in having two inputs and with the second output dependent on the first one. Keep in mind that while the one qubit gates can work with states in superposition, the two qubits gates can eventually work with two entangled particles.

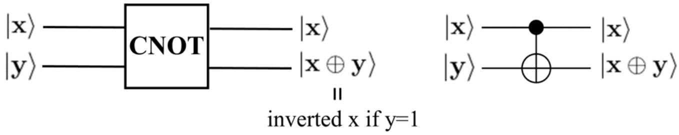


Fig. 52 The two symbols for the controlled NOT quantum gate.

Therefore, if entangled, their input and outputs must be represented by state products. In Fig. 52, the symbol $\oplus$ denotes an operation that gives output x if $y = 0$, or x negated ('flipped') if $y = 1$ (more technically: an addition modulo 2 or a classical XOR operation). The first qubit $|x\rangle$ is left unaltered while the second acquires the value of the first if the second was initially zero, its negation otherwise. This can be summarized in Table 2.

In	Out	In	Out	In	Out	In	Out
0	0	1	1	0	0	1	1
0	0	0	1	1	1	1	0

Table 2 Truth table of the CNOT gate.

A final QLG worthy of mention is the '*Toffoli gate'*, which is a three-qubit CNOT gate version. In fact, its functioning can be summarized by recalling that the first two qubits remain unaltered while the third is inverted only if the first two are one.

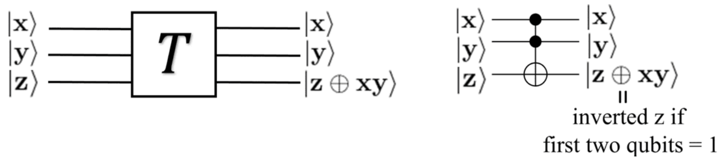


Fig. 53 The quantum Toffoli gate.

Let us now build a very simple quantum circuit with some of these QLG as a proof of concept to clearly demonstrate the advantage of QC compared to its classical counterparts, at least in some specific settings.

One of the most simple and insightful quantum circuits is implemented by the two-qubit version of the 'Deutsch algorithm' (due to the British physicist David Deutsch, who proposed it in 1985). Deutsch's algorithm determines whether a function $f(x)$ is constant or balanced, that is, whether $f(0) = f(1)$ or $f(0) \neq f(1)$, respectively. Given a set of two bits $\{0,1\}$ as entry values of function f, one would have to test the function f on both bits to see if it is constant or balanced. Whereas, as we are going to show, using Deutsch's algorithm, only one query is necessary for a quantum computer. It is the most primitive (not really useful) example of '*quantum query*' but it serves well our didactical purposes and is illustrated in Fig. 54.

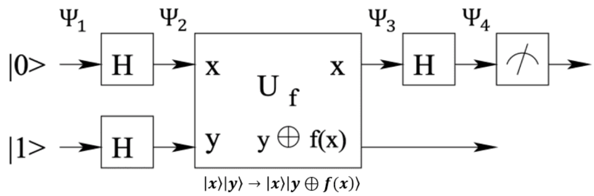


Fig. 54 The Deutsch algorithm.

First, two qubits are prepared in the quantum state: $|\Psi_1\rangle = |0\rangle|1\rangle$. The second quantum computational step superimposes the input vectors, applying a Hadamard transformation that operates on the basis vectors as:

$$|\Psi_2\rangle = H|0\rangle \cdot \mathrm{H}|1\rangle =$$

$$= \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right) \cdot \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) = \frac{1}{2}(|0\rangle|0\rangle - |0\rangle|1\rangle + |1\rangle|0\rangle - |1\rangle|1\rangle).$$

Because the oracle U_f maps $|x\rangle|y\rangle \to |x\rangle|y \oplus f(x)\rangle$, the third computational step, that is, its application to $|\Psi_2\rangle$, returns (noting that, in general, $0 \oplus f(x) = f(x)$):

$$|\Psi_3\rangle = U_f|\Psi_2\rangle =$$

$$= \frac{1}{2}(|0\rangle|f(0)\rangle - |0\rangle|1 \oplus f(0)\rangle + |1\rangle|f(1)\rangle - |1\rangle|1 \oplus f(1)\rangle). \quad \text{Eq. 21}$$

Notice that $|\Psi_3\rangle$ is an entangled quantum state in which f(0) and f(1) enter the computation step simultaneously, in contrast to classical information processes which must evaluate them separately.

Now let us consider the first case in which the function is constant. Then, replacing in the above equation $f(1) with f(0)$, we can simplify $|\Psi_3\rangle$ to:

$$|\Psi_3\rangle = \frac{1}{2}(|0\rangle|f(0)\rangle - |0\rangle|1 \oplus f(0)\rangle + |1\rangle|f(0)\rangle - |1\rangle|1 \oplus f(0)\rangle) =$$

$$= \frac{1}{2}(|0\rangle + |1\rangle)|f(0)\rangle - (|0\rangle + |1\rangle)|1 \oplus f(0)\rangle)$$

$$= \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right)\left(\frac{|f(0)\rangle - |1 \oplus f(0)\rangle}{\sqrt{2}}\right)$$

$$= \frac{1}{\sqrt{2}} H|0\rangle \; (|f(0)\rangle - |1 \oplus f(0)\rangle),$$

with the last passage because of the Hadamard operator acting on the first qubit $|0\rangle$, as we had evidenced in Fig. 50.

Finally, one interferes this output with the H-gate again and this furnishes the result (recall that $H \cdot H^* = I$):

$$|\Psi_4\rangle = \frac{1}{\sqrt{2}}|0\rangle \; (|f(0)\rangle - |1 \oplus f(0)\rangle) .$$

The state of the second qubit is given by a somewhat complicated expression, but this should not concern us. Whereas, the statement on the first is clear: If the function is constant, a measurement on the first qubit (evidenced with the symbol after Ψ_4 in Fig. 54) will show it to be in state $|0\rangle$.

One must now repeat the same calculation, considering the second case in which the function is balanced, that is, set $f(0) \neq f(1)$. To keep things simple, suppose that f is a binary function so that $f(0) = 0$ and $f(1) = 1$. This means that $1 \oplus f(0) = f(1)$ and $1 \oplus f(1) = f(0)$, and replacing this in Eq. 21, again collecting the first and second qubit terms, one gets (take paper and pencil and check):

$$|\Psi_4\rangle = \frac{1}{\sqrt{2}}|1\rangle\,(|f(0)\rangle - |f(1)\rangle),$$

Which tells us that the first qubit is in state $|1\rangle$.

The bottom line is that QC allows us to do calculations on all possible states in one iteration. In fact, by checking the first qubit, Deutsch's algorithm is able to determine whether a function is constant or balanced with only one quantum evaluation, computing f(0) and f(1) simultaneously. This is in contrast to the classical case, in which two steps are necessary. The first qubit is entangled to the other one, but the former does not take part in the computation. Therefore, the first can be probed without causing the state function collapse of the second one. This scheme can, in principle, be extended to a much higher number of bits (a generalized version called '*Deutsch–Jozsa algorithm*').

There exist many other quantum algorithms that have a definite advantage over their classical computation counterpart. Without going into the details, one could mention '*Grover's algorithm*' which, in a style similar to Deutsch's search query algorithm, enables one to find a specific string of qubits within a large database. If a database contains N items, the time to find a specific item required by a QC is proportional to the square root of N. This is in contrast to the classical computer, which needs a time proportional to N. This is an enormous leap if one considers that, for example, for 10^6 items the classical computer must 'look them all up', one by one, whereas the QC needs no more than 1000 computations.

Among other things, half-adder and adder circuits have been realized conceptually and, to some degree, also practically (with all the limitations that we will discuss next – state of the art of 2019) and which would pave the way to more traditional mathematical quantum computation circuits.

3. Quantum cryptography

One of the potentially most useful applications of the simultaneity of quantum superposition, entanglement, and interference could be '*Schor's* *algorithm'*, which enables a QC to factorize integer numbers into their prime factors (say, for instance, $15 = 3 \cdot 5$ or $189 = 3 \cdot 3 \cdot 3 \cdot 7$, etc.). Classical computers can do this easily with small numbers but, for huge numbers, the number of necessary computational steps increases exponentially and, above a certain magnitude, becomes practically prohibitive. On the other hand, using Shor's algorithm, the number of steps is proportional to the size of the number (so-called '*polynomial time'*) and allows for much more efficient computations. Integer factorization is a very important mathematical operation that stands behind modern encryption technologies. All our modern digital economy relies on it. Whenever you use your credit card, all the information is encrypted using integer factorization of huge numbers. This makes the decryption—that is, the attempt to crack the code—computational intractable. For this reason, this type of encryption method has turned out to be quite efficient and is used nowadays in most applications that require high-security data exchange. The coming of QC is bad news in this respect because it might change this state of affairs, being much faster in factoring numbers (recent studies suggest this isn't entirely certain though [22]).

However, the good news is that QC itself might furnish a new system of quantum encryption which could be even more secure than classical cryptography. Quantum cryptography might even be the ultimate unbreakable encryption method, even for QC. If you have followed so far all that we said in the chapters about photon polarization, photons being in a polarization superposition state, photon entanglement, the use of polarizers to check Bell's theorem, and how Alice and Bob can use polarized photons as a means of transmitting information by tilting their polarizers, it should be easy at this point to understand how quantum cryptography works.

An interesting application is to use quantum effects to test whether someone is listening along a communication channel between Alice and Bob. At the same time, this test can be used to exchange a secret public encryption key necessary to encode the message containing the sensitive content.

Suppose Alice would like to send a message to Bob through a quantum communication channel such as that of Fig. 55. It is supposed to be a one-way channel from Alice to Bob. (Of course, Bob can use the same technique in the opposite direction.) To do so, she first sends a sequence of bits to Bob (still not the sensitive message) that has a double function: first, that of testing whether someone is listening along the communication channel and second, to provide Bob with a quantum encryption key. This can be done as follows.

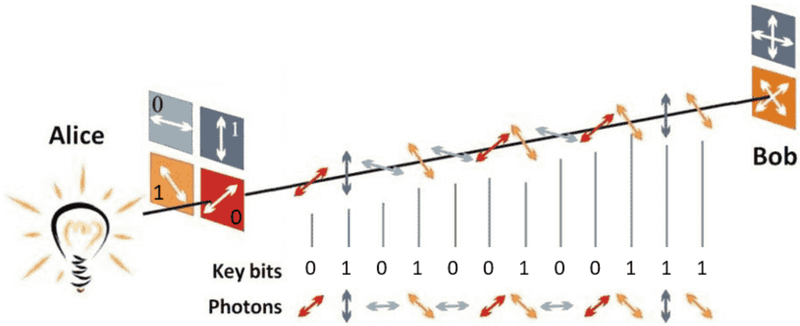


Fig. 55 Polarized photons communication channel.

Alice sends random linearly polarized photons to Bob along each of the possible 45°, 90°, 135°, and 180° orientations. These correspond to two polarization bases. The rectilinear polarization base (photons with 90° and 180° eigenstates) we could label as ⊕, while the diagonal polarization base (photons with 45° and 135° eigenstate) we could label as ⊗. Both Alice and Bob agree to assign the 45° and 90° polarized photons a bit value of 0, and to assign the 135° and 180° polarized photons a bit value of 1. This implies that the choice of the basis still does not tell us anything about the value of the bit.

For example, say Alice sends a 45° polarized photon (0-bit information). If Bob uses a diagonal base polarizer, he has a 100% chance of detecting it (recall Malus' law). And with a second linear polarizer, he could establish, with certainty, whether it is a 1- or 0-bit photon. Say he uses a 45° polarizer: The 0-bit photon goes through for sure, while the 1-bit photon will certainly be blocked (one assumes that Bob and Alice are synchronized and the time of arrival of all photons is known: If no photon is observed Bob knows it must have been blocked).

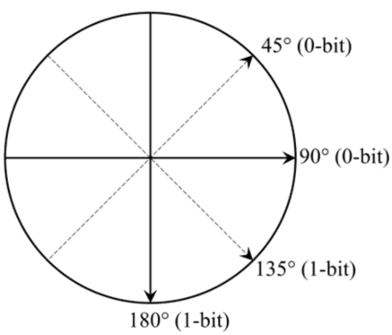


Fig. 56 Photon polarization for the 0- and 1-bit value.
Solid arrows: rectilinear polarizers ⊕. Dashed arrows: diagonal polarizers ⊗.

So, every time Alice and Bob use the same polarization base (and assuming, for the sake of simplicity, that the communication channel hardware is ideal and does not introduce any disturbances), their bits must correspond. That is, whenever Alice and Bob use the diagonal (rectilinear) base polarizer for a 0-bit (1-bit) photon, they must agree on the photon bit value as well.

If, however, Bob uses a rectilinear polarizer, the 45° photon will be set in a superposition of the vertical and horizontal states. Using a linear polarizer again will therefore reveal Alice's original bit value with only a 50% chance. Say Bob uses a 90° polarizer: For half of the cases, the photon makes it through and he will correctly assign it a 0-bit value, whereas if nothing is visible, Bob would wrongly conclude that the original photon's state must have been a 180° 1-bit photon.

Note that this 50% chance is a purely quantum random process. There is no way, not even in principle, to know in advance whether Bob measures a 1 or 0-bit. It is for this unpredictable quantum randomness that quantum cryptography is supposed to be unbreakable.

So, on average, whatever polarization bases are chosen, Bob will decode Alice's photons correctly only for three cases out of four, and about 25 % wrongly. Of course, the same applies to all the other 90°, 135°, and 180° directions Alice could have chosen. The important point to keep in mind is that there is always an inherent quantum error rate that characterizes this communication channel.

This is not a very efficient method of transmitting information. It is a method to produce a quantum random stream of bits with a fixed average error rate detection. Once Alice and Bob compare their data (meeting or communicating through a classical channel), they can verify the error rate. Moreover, they can agree that all those photon matches that turned out to be correct must be taken as the bit sequence that forms the encryption key.

Let us suppose that Alice sends to Bob the following random sequence of bits corresponding to their respective bases.

$\oplus$	$\otimes$	$\oplus$	$\oplus$	$\otimes$	$\otimes$	$\otimes$	$\oplus$	$\otimes$	$\oplus$	$\oplus$	$\otimes$	$\oplus$	$\oplus$	$\oplus$	$\otimes$	$\oplus$	$\oplus$	$\otimes$	$\oplus$
1	0	1	0	0	1	0	0	1	0	0	0	0	1	0	0	1	1	0	1

Bob chooses to randomly measure each single photon along one of these bases and determine its linear polarization. (Well, he 'determines' nothing, as you should know, but 'projects' it to one of the two possible eigenstates!)

$\oplus$	$\oplus$	$\otimes$	$\oplus$	$\otimes$	$\otimes$	$\oplus$	$\otimes$	$\otimes$	$\otimes$	$\oplus$	$\otimes$	$\otimes$	$\oplus$	$\otimes$	$\oplus$	$\otimes$	$\oplus$	$\oplus$	$\oplus$
1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	0	1	1	1	1

Now Bob sends to Alice, through a classical communication channel, the sequence of polarisation bases he used, with the bit values he measured. Alice can then compare the results and retain only those measurements that matched on both sides the polarisation base—that is, also their bit values. If nobody is interfering with the communication and the communication channel is ideal, the binary matches (the gray shaded bits) must be: 1001100111. This digital sequence could be used as the encryption key. (Of course, this was only a simplified example; much longer sequences would be used in practice.)

However, suppose that between Alice and Bob there is Eve, who is secretly listening. She also tries to establish Alice's photon polarization, trying to guess randomly which polarizer base was used to intercept her photons.

Eve will perform the same type of measurements that Bob is supposed to make. If her polarizer base matches Alice's photon polarization (say, the photon is 45° polarized and Eve uses a diagonal base polarizer), Eve will correctly interpret Alice's bit. If a new photon is forwarded with the same polarization orientation to Bob, he won't notice anything in his error statistics. But, if she uses the opposite polarization base that does not match Alice's photon polarization, the photon will again be set in a polarization superposition and Eve will no longer be able to establish, with certainty, the photon's original polarization. She has only a 50% chance of correctly guessing whether it was a 0- or 1-bit photon.

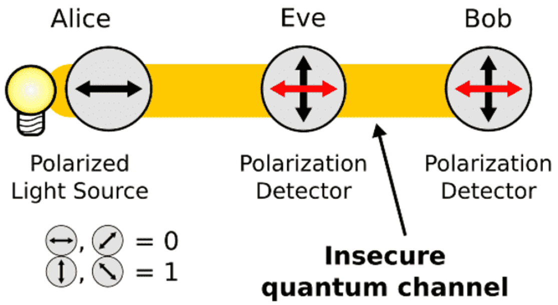


Fig. 57 Principle of quantum cryptography.

But the real problem she will face is that in both cases she will measure the wrong polarisation, since the rectilinear base sets the original 45° photon into a 90° and 180° polarization superposition. This means that Eve can't say whether the original photon was either a 45° or 135° polarized one, and which she set in superposition, or really a 90° or 180° photon that went through the rectilinear polarizer without state change. She is forced to guess or simply forward to Bob a 90° or 180° polarized photon when Alice sent a 45° one, increasing the bit error rate.

This is an obvious fact, after all, because Eve and Bob serially use the same quantum error-prone detection scheme, which can lead only to an accumulation of error. Moreover, as we have already mentioned in the chapter on quantum teleportation, it can be shown that a '*no-cloning theorem*' applies, according to which there is no way, not even in principle, to exactly clone photons or any quantum state. Therefore, an ideal non-disturbing measurement is out of the question. Eve's attempt to listen will inevitably add some amount of noise to the communication. (Here, the microscope analogy for Heisenberg's uncertainty fits.)

Nevertheless, at this stage, Bob still does not notice anything. He may also not notice anything strange in the intensity flux of the photons arriving if each bit is encoded in only one photon. This is because Eve replaces, one by one, each photon that passes by. (If Bob would observe an intensity variation, then he would know for sure that someone was listening in the first place.) However, this time, the match between Alice's and Bob's polarization bases no longer guarantees the match with the bit value. This is because even if Bob uses Alice's diagonal base (that of her 0-bit valued 45° polarized photon), in this case, he receives Eve's 90° or 180° polarized photons and there is only a 50% chance that he will read the former polarization—that is, Alice's correct and original 0-bit value.

Therefore, once Bob has sent his sequence of polarization bases and bits to Alice via the classical communication channel, she will certainly notice Eve's 'eavesdrop'. She will, first of all, observe a deviation in the bit error rate—that is, a higher than 25% mismatch between her sent bits and those that Bob measures. And, most importantly, she will observe that the polarization base match does not correspond to a bit match: a clear sign that someone is listening. Eve might initially have some luck reproducing Alice's values simply by quantum coincidence. But, after n photons, the chance that, despite Eve's quantum eavesdrop, this will still be the case decreases as $\frac{1}{2^n}$: something which will make it quickly a practical impossibility for a sequence larger than a few bits.

Eve might have also hacked the classical communication channel between Alice and Bob. Nevertheless, that would not help her much if she does not have access to Alice's polarization bases. Note that the single photon encoding is a decisive factor. Eve might attempt to listen by picking up only a few photons at a time, inducing only a small deviation in the expected error statistics. However, she will then miss the public encryption key, without which she can't decode the message, even with a prime number factorizing QC, as she doesn't have the number to factorize in the first place.

So, in principle, quantum cryptography seems to be an unbreakable method. Yet, whether that is really the case and whether QC and quantum cryptography will revolutionise future IT and cybersecurity are situations that remain to be seen. There is still no consensus and it remains a matter of heated debate among scientists as to whether quantum encryption is truly the ultimate secret tool and one that is not susceptible to hacking.

4. The long road towards quantum computing

As of 2019, how far is present technology from realizing quantum computation and quantum cryptography?

Several technical limitations must still be overcome. For example, single photon sources and detectors are far from having 100% efficiency. That is, sometimes they produce more than one, and sometimes none in a still uncontrollable fashion. However, the hardware's efficiency limitations are not the primary cause of concern. It is to be expected that these can probably be overcome with more efficient devices resulting from further technological progress.

Some elementary LQG and circuits have already been realized and tested in the laboratory. In principle, it has already been shown that quantum computing works. However, one technically almost insurmountable obstacle remains which is much more worrisome, making it still unclear whether QC will be practically available soon and whether it is even feasible at all: the tendency of entangled qubits to decohere! The tiniest disturbance leads the entangled particles to collapse back into their '*pure state*', the state corresponding to an isolated quantum particle, in contrast to their multi-particle *'mixed stated'* (see chapter IV 5). For this reason, quantum particles, and the gates and circuits they represent, must be perfectly protected from the environment and from the surrounding radiation in an almost absolute zero vacuum at millikelvin temperatures. Otherwise, the stuff making up a QC gets too noisy and the qubits are flipped and randomized, making any computation processes useless. Due to the background noise, the actually implemented quantum algorithms perform correctly in only 70% of cases. This is a still an unacceptable error rate that makes it clear that, if there isn't a major breakthrough, QC might still be far away.

Decoherence and noise also make it extremely difficult to build QC with only a few qubits. The main problem is that the noise in quantum circuits increases exponentially with the scaling of the circuits and the number of qubits. IBM once predicted that the time of '*quantum supremacy'*, that is, when QC will be able to perform computations which not even the fastest classical supercomputer is capable of, will come once a 50-or-more-qubits QC is built. However, when IBM and later Google did, in fact, build such a QC, it turned out to be affected by too much noise and it was far from being able to even nearly compete with standard machines. In principle, it would be possible to apply error correction methods to the computing system, but this also implies that several entangled particles must represent a single qubit. This, in turn, implies that an even more sophisticated technology will be needed to contain decoherence. If and how quantum supremacy is establishes remains a controversial issue. The time of quantum supremacy seems to not be just around the corner.

Scepticism regarding QC is growing. The most prominent figure among QC sceptics is Gil Kalai, an Israeli mathematician, who claims that quantum computing is an impossibility. This is not because of technical limitations but, according to Kalai, because it is impossible, even in principle, to get the noise down, regardless of the technology one uses, due to the limitations imposed by the fundamental theorems of computation. Not all scientists share his view. Kalai's mathematical proof is not disputed but most regard it as a formal oversimplification of the real, practical implementation of QC. Who is right? Only time will tell.

These are the facts surrounding QC and that we have presented in this chapter. However, the author can't refrain from adding his own two cents. Elsewhere, I outlined in detail my critical stance towards the multi-billion-dollar investments in mammoth science projects to the detriment of smaller and more creative science projects. [23] It is now about a couple of decades since worldwide intense research was launched into quantum computing. Progress has been made from a theoretical point of view but, at the practical level, the advance has not met the expectations. The technological development pace of QC, compared to that of digital computers, is disappointing. The first digital computers appeared after WWII, and within 20 years commercially available 16-bit minicomputers were in existence. Additionally, the transistor and integrated circuits became commonplace. Most of the evolution of classical computers was not due to a concerted big global science project but, on the contrary, came initially from secret military research and later was driven by commercial impetus in hardware development. Every year reported a breakthrough. However, the same could not be said about the last 20 years of R&D in QC. Despite research centers worldwide being funded with billions of dollars in investments, and the current involvement of giants like IBM, Intel, Microsoft, and Google, we are still far from having a single working QC model capable of doing what the first classical computers could do in the mid-1940s. Of course, comparisons are always subjective and debatable, but the history of science repeatedly shows that scientific and technological progress cannot be predicted and, even less, guided in advance by wishful thinking. In fact, as we have pointed out several times, it is the research in QC that led to the technological breakthroughs which allowed for the realisation of what were once considered only thought experiments and which led us to so many new insights in the foundations of QP. However, interestingly enough, this is exactly the opposite result of the original intent of the pragmatist approach towards quantum computing—namely, the practical realisation of new machines and their potential applications and commercialisation, such as in cryptography or artificial intelligence. The positive evolution in fundamental science, as well as the development of new technological devices that changed the world, have always required free spirits who are able to express their creativity and genius independently. It does not come from a science that requires a permanent media circus, without which it would run out of funds, with such curious side effects as the proliferation of predatory journals and fake international conferences. In the context of a hierarchically structured managerial market-driven enterprise and the purely utilitarian approach of big science initiatives based on a '*publish or perish*' academy, almost all the free spirts have been killed off and brought to the verge of extinction. And what if Kalai's conjecture turns out to be correct? That would be the final nail in the coffin: an intellectual, scientific, and financial catastrophe for all those who worked so hard in the last decades on the realisation of QC. Will this be another '*nightmare scenario*', as we already hinted at with the case of superstring theory? Hopefully not, but the signs are all there.

5. Classical information theory

The recent research on QC took up and revived another subject that was long pursued by some pioneering physicists – namely, the connections between information theory and QP. In general, information theory is a mathematical theory that arose in the 1950s with the advent of modern classical computer technologies. In the collective consciousness of mathematicians, IT engineers, and physicists, Claude Shannon, an American mathematician and engineer who first introduced the notion of '*information entropy*', is considered the father who provided the foundational pillars of this theoretical science. Soon it became clear that it could be extended to QP, as the wavefunction (or the state vector) and all the quantum algebra related to it can be considered an informational theoretic language. In 1990, J. A. Wheeler proposed an "*it from bit*" doctrine according to which information – and not matter, energy or space-time – should be regarded as fundamental to the physical universe. According to Wheeler, all things physical are information-theoretic in origin. However, it was in the last couple of decades, when scientists pursued the construction of QC, that this science received a new impulse. Information theory also found its way to modern QG theories. There is a large consensus that information must indeed be something fundamental to our existence and that, whatever the fate of QC will be, it might become one of the key ingredients necessary for a general QG theory that unifies the four fundamental forces of Nature. It is, therefore, worthwhile to take a look at this fascinating and promising subject by reviewing some of its concepts that the reader will most likely encounter frequently in the future.

The history of information theory can be traced back to the birth of thermodynamics, especially the development of that concept that is nowadays in many scientists' mouths, which is entropy, and that later paved the way to statistical physics.

The thermodynamic fundamental quantities, so-called '*state variables*', describe the state of a thermodynamic macroscopic system. The most common ones are pressure, volume, temperature, and mass. With these, one can, for instance, describe the state of an ideal gas. Recall (see the chapter on the black body radiation in Vol. I) that we speak of a thermodynamic state of equilibrium if the state parameters do not change in time – that is, there are changes in neither the temperature nor any mechanical or thermal quantity. Simply imagine a gas in a room at constant temperature. This can be contrasted with a *'thermodynamic transformation'*, which is defined as a thermal evolution of a system from one equilibrium state to another. There can be reversible and non-reversible transformations. A *'reversible transformation'* is characterized by its ability to be reversed in such a manner that the system can be brought back to its starting point without any sort of dissipation. These are only ideal transformations which cannot exist in practice but are good approximations for several cases of interest. An irreversible transformation is obviously a transformation which is not reversible, and which is of the kind we always observe in reality.

The first to introduce the notion of entropy was Sadi Carnot, a French engineer who, in 1824, published a book in which groundbreaking theoretical discoveries went unnoticed for a long time, but which is nowadays considered one of the foundational pillars of modern thermodynamics. Carnot was busy developing a theory for steam engines and showed, among several other things, that some abstract – and until-then undefined – quantity is conserved during an *'isothermal'* (constant temperature) reversible transformation. He realized that for any reversible thermodynamic process, the factor between the heat exchange dQ and the temperature T is always equal. And when, in physics, something invariant appears, it usually signals an important quantity. Carnot generalized it to an expression that considers an exchange of infinitesimal quantities of heat as:

$$dS \geq \frac{dQ}{T}. \quad \textit{Eq. 22}$$

The reason for the inequality will be explained soon. Later, the German physicist Rudolph Clausius called this quantity *'entropy'* (from the Greek word ευτροπια, which means 'change', 'variation, or 'evolution'). It is nowadays conventional to label it with a capital letter S (perhaps standing for 'State' but, curiously, nobody knows for sure). Entropy measures the change that a thermodynamic system experiences in a reversible or irreversible process. If it is a reversible cycle (that is, a transformation that brings the system back to its initial conditions without dissipating heat), the equality holds and we have no entropy change. If, however, the cycle is irreversible (that is, the system turns back to the initial conditions but has dissipated some heat, as must happen for every realistic thermodynamic system), entropy always increases. From this, it follows that, in reality, dS is always > 0 – that is, some heat will always be left here or there during a transformation. This is one way to express the famous *'second law of thermodynamics'*. There are other, equivalent versions of this law, such as simply stating that in a cyclic process, the entropy in an isolated system never decreases.

Now, you may have heard about entropy as a measure of order and disorder and be wondering how this idea connects to the above classical definition. Carnot, Clausius, and all physicists until 1877, when Boltzmann came up with a statistical interpretation, never used entropy as a measure of disorder. In fact, first-year undergraduate courses on thermodynamics might not even mention entropy in such terms. In physics, entropy is, first of all, a state function of a thermodynamic system which does not necessarily need to make any reference to concepts like order or disorder.

To understand the statistical mechanical point of view of the meaning of entropy, it might be interesting to recall that there is also a somewhat less popular but still meaningful third law of thermodynamics (also called *Nernst's law*) which states that the entropy of an equilibrium system vanishes when the temperature T approaches the limit of absolute zero: $lim_{T\to 0}$ S $= 0$. This suggests that when the Brownian motion of a body is turned off (that is, it is completely frozen and can attain only one state configuration, called *'microstate'* or *'degree of freedom'*), then the entropy vanishes. If, instead, a body has some non-zero temperature, it can attain a huge number of microstates. To better understand the physical meaning of microstates and how these relate to the notion of entropy, let us first make a macroscopic analogy.

As an example, imagine the number of all possible states a dice can be in after it is thrown. (Follow the reasoning by use of Fig. 58.) When thrown or rolled, it comes to rest in only one of the possible six states W: 1 to 6. We have probability $p = \frac{1}{6}$ to get, for example, 6.

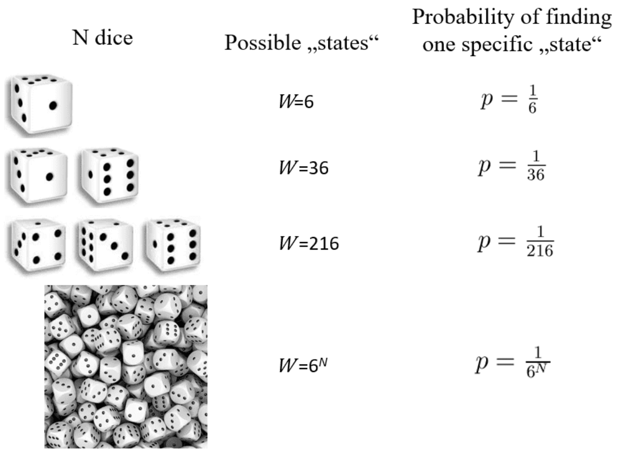


Fig. 58 Analogy between 'dice-states' and microstates.

If we throw two dice, there are $W = 6 \times 6 = 36$ possible outcomes – i.e., the number of possible states in which the system of two dice can be found. Then the probability that the two dice will both show the face with number six is $p = \frac{1}{36}$. For three dice, we already have $W = 6^3 = 216$ possible outcomes, or states, and a probability of $p = \frac{1}{216}$ to find all three dice showing the number six. In general, for a large collection of N dice, the number of possible states that the system of dice can acquire can become an incredibly huge number and scale as $W = 6^N$and $p = \frac{1}{6^N}$.

Something similar happens with real gasses in thermodynamics. If, instead of conceiving of dice we think of the huge collection of minuscule atoms or molecules of which a gas is made and label as 'state' the molecules' microscopical energy state, or position state, or rotational and vibrational state, etc., then we can speak of all the possible microstates a gas can be in without showing an appreciable change at a macroscopic scale. The system can attain a huge number of possible states, also called a *'statistical ensemble'* (ensemble: from the French meaning 'a collection or a set of things'), which, however, at a macroscopic scale cannot be distinguished from each other. Imagine, for instance, a gas at room temperature: Its molecules continuously zigzag in every direction and we would observe, at very small scales, a sort of *'molecular chaos'*. However, at our human or greater scales, if temperature and all the other thermodynamic variables are kept constant, we don't observe any change in that gas. It is like shaking a box full of dice: They will show a combination of possible states which changes continuously but, from the outside, you see only the box containing it, without noticing the internal change.

Let us see what the order of magnitude of the number of particles is that statistical physics deals with. Take, for instance, about 18 g of water. This is approximately one mole of H_2O molecules. If you studied chemistry in high school, you know that a mole of any substance contains something like $6{,}022 \times 10^{23}$ molecules (or atoms). This is the *'Avogadro number'* and it is enormous! Even with the most powerful supercomputer, it is impossible to calculate and simulate the behavior of every molecule in such a gigantic set. Therefore, scientists are forced to resort to statistical approaches, which means we can describe only the average properties of the collection of all these particles as a whole and ascribe to it some macroscopic values – that is, state variables such as a volume, a pressure, a temperature, etc. It is possible to statistically study some mean value, such as the mean energy one must expect for one particle averaging over that of all the particles. Statistics allows us to compute the probability that one single particle is in a specific state, or the average of some property; to know precisely in which of the possible gazillions of microstates a system is in at a specific instant is completely out of the question.

In fact, let us see, once the thermodynamic state variables of a system are known, how many possible microstates it can acquire. It turns out that the number of possible configurations that a system can be in is proportional to Carnot's and Clausius' entropy. Ludwig Boltzmann first made the connection and could show that entropy is related to the number of equiprobable microstates W as:

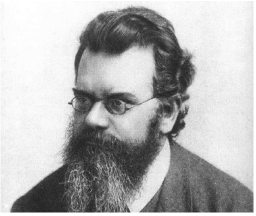


Fig. 59 Ludwig Boltzmann (1844-1906).

$$S = k_B \cdot log\, W, \quad \textit{Eq. 23}$$

with $k_B = 1.38 \times 10^{-23} \frac{J}{K}$ the *'Boltzmann constant'*, having the unit dimension of Joule/Kelvin (energy/temperature) and appearing throughout thermodynamics, especially in describing the kinetic energy of particles. This is Boltzmann's famous formula that you will find engraved on his tombstone. As an example, a minute entropy increase of $\Delta S = 10^{-14} \frac{J}{K}$ leads to an increase of $\Delta W = e^{\Delta S/k_B} = 2.026 \times 10^{314706146}$ possible microstates.

This is an incredibly huge number which goes beyond our wildest imagination. And if you consider that the order of magnitude of a more realistic entropy change of a few moles of a gas changing temperature by about 20 degrees amounts to an entropy change of 10, you will understand why your computer will give you the overflow message error when trying to calculate the possible number of degrees of freedom. That's also why Boltzmann used the logarithm function which has the mathematical property of scaling down awkwardly huge numbers and reducing them to humanly understandable quantities. (See Appendix A Ib.) Moreover, if you consider that, if the number of microstates of two systems is respectively W_1 and W_2, the number of microstates of the whole system containing both will be its product $W = W_1 \cdot W_2$. (Just think of it with simple examples such as two dice.) Also, for this reason, the logarithm turns out to be very useful because it is the only additive elementary function, that is, $\log a \cdot \log b = \log(a + b)$, which tells us that the sum of the entropies of each subsystem must furnish the entropy of the whole system, just as the mass of an object must be the sum of the masses of its constituents. This statement sounds self-evident, doesn't it? But as we shall soon see, as usual, Nature will teach us another lesson.

Therefore, in Boltzmann's interpretation, what entropy tells us about is the incredibly huge amount of ensembles of *possible* microstates in which a thermodynamic system can be. However, we cannot distinguish them from each other at a macroscopic scale. Because every microstate is macroscopically equivalent at our human scale, and because of the large number of particles involved, we always see the same system, even if it undergoes a continuous microscopic change from microstate to microstate without our being aware of it.

This can be reframed in other conceptual terms. Namely, the statistical mechanical interpretation of entropy also suggests the measure of our ignorance in knowing precisely in which microstate the system is in. We can only say that it must have one of the W possible ones but we can't say with certainty which one. And this is the limit that statistics has to deal with. It is in this sense that entropy could be used to characterize the 'disorder' of the system, as it tells us something about the molecular chaos occurring in it as well as about the ignorance of our minds in relation to the observed system.

Later, J. W. Gibbs, an American engineer (who, by the way, coined the term *'statistical mechanics'*) further developed Boltzmann's formula, rewriting it in terms of probabilities. Say that p_i is the probability that a thermodynamic system is in the i-th microstate (and by now you should understand why we are talking about extremely tiny probabilities). Then the entropy written in terms of all the possible microstates a system can have is the *'Boltzmann-Gibbs entropy'*:

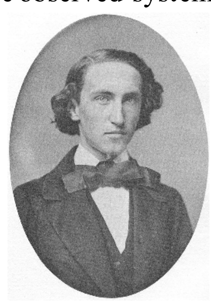


Fig. 60 Josiah Willard Gibbs (1839–1903).

$$S = -k_B \sum_i p_i \cdot \log p_i \quad \textit{Eq. 24}$$

This is essentially the same expression as Boltzmann's original entropy in terms of microstates but expresses it in a different way, over a sum (a huge sum!) over all the possible microstates, considering also the probability that each microstate is realized. (If all the W states are equally probable, then $p_i = \frac{1}{W}$ and one obtains, again, the Boltzmann entropy of Eq. 23. See also end of Appendix A Ib) The negative sign reminds us that it is about an ignorance, a statistical (still classical!) uncertainty, or about a lack of information. Another way to think of entropy is as a 'measure of diversity'. In this way, physics understands the notion of 'information' as a physical property of reality.

That's how scientists developed the fields of thermodynamics and statistical mechanics, which satisfyingly describe the thermal properties of bodies in terms of macroscopic parameters without the need to know everything about each of its microscopic atomic or molecular constituents. In this technical and historical frame, the concept of entropy and information was born.

The introduction of probability to physics can successfully describe physical states and properties without any mention of QP. In classical Newtonian mechanics, probability is introduced not because of an objective indeterminism as in QP but because of our ignorance and our minds' inability to keep every particle of a system under control and to know every one of its instantaneous microstates. Maybe Laplace would have been disappointed. His ideal of a precise determinism, at least in its strict version, which wanted to know everything about the state of the universe, can no longer be considered practically feasible. Nevertheless, the great success of statistical mechanics is just that it could show that it is not at all necessary to know all the microscopic details. Physics, probability theory, and statistics describe well gases and their thermodynamic properties. His deterministic 'clockwork universe' still holds. Through a statistical and molecular approach, a strictly reductionist point of view of nature prevailed.

Although there is no scientific unanimously-accepted definition, we might say that by a *'reductionist'* approach in physics, one usually intends that physical phenomena can be explained in terms of the interaction of lower-level entities, such as particles, atoms, or molecules. Every large-scale aggregate in comparison to these elementary entities like particles and atoms, from a human being to a Galaxy, is considered nothing other than the sum of these constituents. This is the decisive point that distinguishes reductionist thinking from non-reductionist thinking and does not need any further explanation: Everything from the dynamics of a subnuclear particle to a Galaxy, such as the philosophical deeper questions about the birth of the universe, an eventual purpose, and final causes, can be explained away by analysing the interaction of these constituents. Implicitly, what this assumes is that the whole is nothing other than the sum of the parts. In biology and psychology, this goes far in the attempt to explain life as a function of trillions of tiny cells, and consciousness as an epiphenomenon of these interactions. However, as we will see in the chapter on quantum consciousness, this was – and remains – only a hypothesis and a debate that is far from being settled. Furthermore, you might recall, from our discussion in Vol. I on the Schrödinger's cat paradox and the measurement problem, how there seems to be a gap between the microworld and the macroworld. The world we experience cannot be derived from QM alone, at least not in its present version. The macroworld cannot be reduced to the microworld. At any rate, by resorting only to Newtonian mechanics, statistical physics could successfully derive many properties of inanimate matter.

The next step in the history of information theory came from Claude Shannon, who in 1948 discovered how Eq. 24 can be related to modern information theory. Shannon's achievement was that, for the first time, a rigorous mathematical definition for a quantitative measure of information was created.

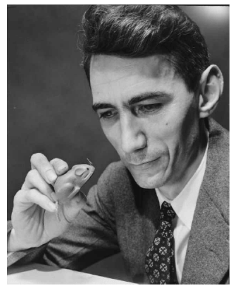


Fig. 61 Claude Shannon (1916-2001).

Shannon was inspired by the previous works of Ralph Hartley, another electrical engineer, who considered digital information stored in memory registers to be made of 'cells' – that is, bits, which can have only two physical states, namely 1 or 0. As we already discussed in the chapter on QC, a register of N bits can have $W = 2^N$ distinguishable states. Hartley, in analogy to Boltzmann's and Gibb's approach to entropy, defined the *'capacity'* C of a register to store information as a quantity proportional to the number of possible different states it can attain: $C = log_2 W$ (the basis 2 of the logarithm function because of the two on-off states of each bit).

However, Shannon was still not satisfied with this definition because, while it said something about the storage capacity of a memory device, it did not say anything about its information content. To clarify the difference, consider the following binary message: "11111111101111111111". This is a message that one can store in a much more efficient way, for example, by agreeing that "zero is the tenth digit and all others can be discarded because we already know that they are one". This reasoning uses the *'surprise factor'* in the sense that the unit value is not surprising, though the zero digit is. Then simply write the number 10 in digital format – namely, "1010". Then the 20-bit message can be compressed into a four-bit one. This logically implies that because a method has been discovered to compress the information content of a 20-bit string of data into a four-bit-long string, there might exist an even more efficient compression algorithm that can do the same with even fewer bits. Shannon, therefore, defined information as a *'surprise measure'* or, more precisely, the minimum capacity a storage device requires to be able to store a message. In principle, the number of bits representing the maximally compressed signal could be taken as the measure of the information content of that signal. Unfortunately, it was not clear how to determine this maximal compression and it turned out much later, about the turn of the millennium, that there is no way, not even in principle, to determine this minimum capacity – that is, the maximum compression to which a signal can be subjected. Gregory Chaitin, an Argentine-American mathematician and computer scientist, was able to prove, by studying the mathematical properties of complex systems via a field known as *'algorithmic information theory'*, that an *'information-theoretic incompleteness theorem'* holds, which expressively disallows the knowledge of what the computationally incompressible string for a specific information content is.

Shannon, however, approached the problem by changing from a deterministic description to a probabilistic one and defined, as the measure of information contained in a sequence of N bits,

$$S = -\sum_{i=1}^{N} p_i \cdot log_2\, p_i$$

where p_i is the probability that the i-th bit takes the specific state 1 or 0. Its main difference from the Boltzmann-Gibbs entropy is that there is no Boltzmann constant and the logarithm is in base 2. For instance, in the example binary message above, since on 20 digits 19 where a 1-bit the probability of obtaining a one is 19/20 and the probability of obtaining a zero bit is 1/20, which translates into an information content:

$$S = -19 \times \left(\frac{19}{20} \cdot log_2 \frac{19}{20}\right) - 1 \times \left(\frac{1}{20} \cdot log_2 \frac{1}{20}\right) \approx 1.56 \text{ bits.}$$

Which means that the above 20-bit sequence has an informational content less than a couple of bits or that, at least in principle, a two-bit digit sequence could also encode the 20-bit one. And if we consider a completely random message of N bits, that is, insert for $p_i = 1/N$, then Shannon's entropy results in:

$$S = -\sum_{i=1}^{N} \frac{1}{N} \cdot log_2 \frac{1}{N} = log_2 N, \qquad \textit{Eq. 25}$$

Which tells us that for a truly random sequence of values, the full extent of the register's bits must be used.

It might be interesting to mention how one could also take this as a definition of randomness. A signal is considered perfectly random if its sequence of bits cannot be compressed into a shorter message. If a sequence of bits is incompressible, then it is random, by definition. However, because it is impossible to know whether a string of bits is incompressible, one will never know if it is truly random – a strange fact that makes the reasoning on the concept of randomness logically circular. This has philosophical implications that we will take up in a separate treatise on science, consciousness, and reality.

So far, we have applied information theory to classical systems of Newtonian mechanics. It is time to see how it can also be applied to QP and how that led to profound insights and the fast-developing field of research that is quantum information theory.

6. Quantum information theory

So, what does information theory have to do with QP? To fix the ideas, consider that, if a coin is balanced, the probability of heads or tails showing up is obviously 50%. Shannon information entropy is then:

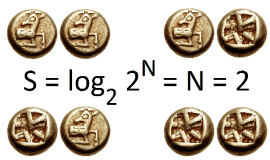


Fig. 62 Two coins have two bits entropy, or two bits storage capacity.

$$S = -\left(\frac{1}{2}\, log_2 \frac{1}{2} + \frac{1}{2}\, log_2 \frac{1}{2}\right) = 1 \text{ bit.}$$

Which states the obvious fact that the system we call 'a coin' can store one bit or, equivalently, it can attain two distinguishable physical states (as $S = log_2 W$, here this implies $W = 2^1 = 2$). Similarly, tossing a two-coin system (N=2) must have two-bit entropy or two-bit storage capacity because it can be in four possible states (heads-heads, heads-tails, tails-heads, or tails-tails), that is, $W = 4 = 2^2 \rightarrow S = 2$.

Let us now compare this with a quantum particle with two spin states. Consider it first prepared in an eigenstate – that is, in a definite state. Recall that once a particle that has been prepared is in its eigenstate, it will show up with the same measured value for every measurement – that is, with the same eigenvalue. Say an electron prepared in eigenstate always shows up with certainty in an up-spin ($|\Psi\rangle = |+\rangle, p = 1$) and this is independent of the number of measurements we will make. How many states can such an electron be in? Only one, as it has been prepared to be in that – and only that – state. What is its associated information entropy? Zero, because if $W = 1$, then $S = log_2\, 1 = 0$. In other words, it has no entropy and no information content (we have no 'ignorance') because there is no 'surprise'. The state of such particles is already known in advance and is always the same. So far, so good.

Now, consider one electron in a spin superposition state, which, as we know all too well, is described by the single state vector (omitting the axis labels) as $|\Psi\rangle = \frac{|+\rangle \pm |-\rangle}{\sqrt{2}}$. The question, again, is: How many states are these? We know that, according to QP, superposition states must be interpreted with a logical AND of being in both eigenstates at the same time. Moreover, due to Schrödinger's equation linearity, if $|+\rangle$ is one solution and $|-\rangle$ is another solution, then $|+\rangle + |-\rangle$ is also one solution—that is, it still represents one single state. Therefore, there is one and only one state, which implies, again, $W = 1$ and $S = log_2 1 = 0$. The particle in quantum superposition has zero entropy.

With this, the analogy with the classical entropy breaks down. While a one-coin system having two discernable states (heads/tails, S=1), the one-particle system in quantum superposition, when considered in isolation, having two states as well (spin-up/down), can be in only one single quantum state (S=0) that, however, once measured, leads to only two (anti-correlated) outcomes. Or, to see it from another perspective, there is a dichotomy between what we measure and what the quantum state is before the measurement. If there were only one single state before the measurement, how could it be that, once we take the measurement, we get two different outcomes?

To further investigate this state of affairs, let us extend this to a two-electron system picked up from a random source and which will therefore show up with two random spin-eigenstates – that is, one is dealing with a statistical ensemble of independent systems given by the state vectors and its probabilities as: $\{(|\Psi\rangle_1 = |+\rangle, p = 0.5), (|\Psi\rangle_2 = |-\rangle, p = 0.5)\}$. Then, in a perfect analogy to the two-coin system, it could be found in the four possible microstates: $|+\rangle|+\rangle$, $|+\rangle|-\rangle$, $|-\rangle|+\rangle$, $|-\rangle|-\rangle$ with equal probability, and S=2. Again, so far, so good.

But what if the two electrons are entangled? We know that the composed system must be described by the single state vector $|\Psi\rangle = \frac{|+\rangle_A|-\rangle_B \pm |-\rangle_A|+\rangle_B}{\sqrt{2}}$. How many states describe this state vector? According to QP, of course, only one, which implies again that the entropy is zero, as if there could be no 'surprise' factor. And yet, when Alice and Bob measure their particles, they will always get one or the other outcome with 50% probability. This is another example of how the entropy concept in QM is somewhat different from that in CP.

One can also extend this to a quantum system of many particles, such as a Bose Einstein Condensates (BEC). As we will see in chapter V.1, a BEC is a quantum gas of many particles at almost absolute zero temperature. Even if it is made up of thousands of non-interacting bosons, it is a system in a single quantum state and, therefore, has zero entropy! However, with only a slight perturbation, the BEC would decohere into thousands of individualized boson-atoms, each with its possible energy level (which brings us back to the great challenge that engineers have to deal with in building quantum computers, as outlined in chapter IV.4). This is what makes the difference between a classical gas and a quantum gas. Despite our mental projection of the letter as a material substance composed of many subunits, as long as it is in a coherent state, its number of possible microstates is only one.

This also leads to the conclusion that, as long as a system remains in a state of quantum superposition and/or entanglement among its subparts, its entropy is less than the sum of the entropies of its subsystems in isolation. It is yet another aspect of QP that speaks against a classical reductionist approach to Nature, which assumes that everything can be described as the sum of its subparts.

Therefore, the fundamental difference between a system of entangled particles or in a superposition state and a statistical ensemble of different and independent quantum states is reflected in the entropy state function. The former quantum systems are said to be in a *'pure state'*, whereas the latter are in a *'mixed state'*.

Pure states have zero entropy, which means that there is no uncertainty regarding the quantum state of the system and which can be described by a single state vector (which, however, might eventually differ for its phase factor), such as in the case of state superposition or particle entanglement, which are **not** a probabilistic mixture of other pure states but, rather, are a unique and undivided quantum state per se.

Mixed states are, instead, characterized by a non-zero entropy, which means that there is uncertainty over the quantum state of the system, and which must be described by more than one state vector, such as a statistical ensemble of independent (sub-)systems, with each (sub-)system being in a definite pure state. The mixed state concept reflects our classical understanding of a system made of subparts, however, it is described by a distribution of physically indistinguishable pure states, in the sense that one does not know which of the many possible states the system actually realizes, just as in classical statistical mechanics, like the Boltzmann-Gibbs gas interpretation with its huge number of possible microstates. Alternatively, a mixed state can represent uncertainty as to which pure state the system has been prepared. As an example of 'noisy preparation', think of unpolarized light, which is a statistical ensemble of randomly polarized photons, though each in a definite state (50% of $|H\rangle$ horizontal and 50% of $|V\rangle$ vertical eigenstates photons). Note how, according to this latter case, a single particle can be considered to be in a mixed state, and a mixed state does not necessarily refer to a collection of particles, as we have also seen in the above examples.

Therefore, the main conceptual difference between a mixed state and a pure state is that, in a mixed state, the uncertainty rises due to a classical statistical uncertainty (that is, because of our lack of knowledge about the real objective state of the system or of its sub-systems), whereas the uncertainty on a pure state is due purely to quantum randomness in the sense that there is only a single quantum state and that there is nothing else to be known (therefore, S=0), though the measurement of the outcome is nevertheless undetermined. In the language of decoherence, one might say that a system of many particles can be in a pure state if there is a quantum phase relationship between its constituents, whereas it becomes a mixed state when this relationship is lost due to decoherence (as the loss of the constant phase relation between particles is ultimately the definition of decoherence). Another (weird enough) way to express this is that, if a total system is in a pure state, its subsystems are in a mixed state.

Having clarified the conceptual foundations, we can go a step further to a more formal representation. The following part and chapter IV.7 are a more complex reading that relies on Dirac's notation formalism; please refer to the chapter on the state vector and Dirac’s notation in Vol. I. We added it here for the sake of completeness for the advanced reader but if you don't feel comfortable with it, don't worry; it is a self-contained part that can be skipped.

A pure state is represented by a state vector of unit length. It can be described by a single ket-vector on a Hilbert space. A mixed state is described by the so-called *‘density matrix’* or *‘density operator’*, usually labeled with the Greek letter ρ (“rho”), and which is a sort of extension of the state vector (or wavefunction) from a single-particle state vector to a many-particle and/or many-state formalism, where the ensemble is given by a probability distribution of states in which the particles can be found. Consider a probabilistic mixture of pure states with each of its probabilities to be in that state: $\{(|\Psi_1\rangle, p_1), (|\Psi_2\rangle, p_2), \ldots (|\Psi_M\rangle, p_M), (|\Psi_{M+1}\rangle, p_{M+1}),$ $\ldots (|\Psi_N\rangle, p_N)\}$. This is a statistical ensemble—that is, a set containing the probabilities associated with the particle's eigenvector space, which means the eigenstates $|e_1\rangle, \ldots, |e_M\rangle$ (again, eigenstates are pure states) and also all the pure states $|\Psi_{M+1}\rangle, \ldots, |\Psi_N\rangle$ resulting from its combination (via superposition or entanglement). All these together form an N-dimensional Hilberst space spanned by the N eigenvectors. Then, in mathematical terms, the density operator is:

$$\rho = \sum_k p_k |\Psi_k\rangle\langle\Psi_k| = \sum_k p_k \ \rho_k(|\Psi_k\rangle) \ , (k = 1,\ldots,N) \quad Eq.\ 26$$

where $|\Psi_k\rangle\langle\Psi_k|$ is, in Dirac notation, a *'projection operator'*. This is an abstract object which you will always find throughout the quantum literature because it formally represents the act of measurement in QM. Measurements described with the density matrix are a more general way to describe the collapse of the wavefunction and decoherence processes.

Let us figure out what this means. Consider first the left-hand side of Eq. 26. A measurement on a quantum system in a generic state Ψ is a 'projection' in the Hilbert space (spanned by the eigenvectors representing all the possible outcome states) representing the outcome of one of these, say, the k-th eigenvector. That is, given the state vector $|\Psi\rangle$, the projection operator $P = |\Psi_k\rangle\langle\Psi_k|$ 'collapses' it into the eigenvector $|e_k\rangle$ with probability p_k, as illustrated in Fig. 63. For example, the usual pure state of spin superposition $|\Psi\rangle = \frac{|+\rangle \pm |-\rangle}{\sqrt{2}}$ is projected into the eigenstate $|+\rangle$ or $|-\rangle$ with probability $p_1 = p_2 = 1/2$. The density operator of Eq. 26 left represents the sum over all the possible projections.

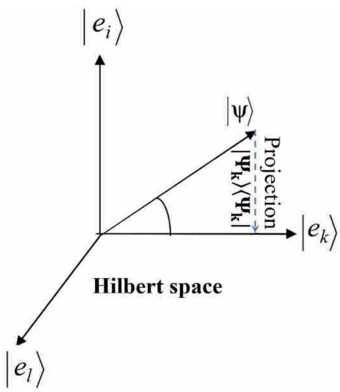


Fig. 63 A measurement represented as a projection operator.

Then, the right-hand side of Eq. 26 expresses the fact that the k-th projection operator is itself a matrix – that is, the density operator ρ is the result of a weighted sum of matrixes:

$$\rho_k(|\Psi_k\rangle) = \begin{pmatrix} \rho_{11}(|\Psi_k\rangle) & \cdots & \rho_{1N}(|\Psi_k\rangle) \\ \vdots & \ddots & \vdots \\ \rho_{1N}(|\Psi_k\rangle) & \cdots & \rho_{NN}(|\Psi_k\rangle) \end{pmatrix}; \quad (i, j=1,\ 2,\ \ldots\ N),$$

with coefficients

$$\rho_{ij}(|\Psi_k\rangle) = \langle e_i|\Psi_k\rangle\langle\Psi_k|e_j\rangle \quad Eq.\ 27$$

The interested reader who would like a more advanced formal understanding can see Appendix A Id.

To see what this means intuitively, recall how we defined the probability coefficients with Dirac notation. (See the section on the state vector and the Schrödinger equation in Vol. I.) A measurement on a system will project the state vector $|\Psi\rangle$ onto one of the possible eigenvectors $|e_i\rangle$ with the probability coefficients $p_i = |c_i|^2$ defined as:

$$|c_i|^2 = |\langle e_i|\psi\rangle|^2 = \langle e_i|\Psi\rangle\langle\Psi|e_i\rangle. \quad \textit{Eq. 28}$$

Eq. 27 is an extension of Eq. 28. In fact, if one restricts, at the diagonal elements of the density matrix, the cases i=j, then Eq. 27 boils down to Eq. 28: $\rho_{ii}(|\Psi_k\rangle) = \langle e_i|\Psi_k\rangle\langle\Psi_k|e_i\rangle = |\langle e_i|\Psi_k\rangle|^2 = |c_{ik}|^2$ (i, k = 1, 2, … , N);

That is, the density matrix is an extension of the modulus squared probability coefficients that, as we know well, determine the probability of a measurement outcome furnishing the i-th eigenstate. Here, it is the probability of finding the system being in the i-th eigenstate of the k-th pure state. Or, in other words, two types of averaging are used: one for the quantum entanglement or superposition of the pure states and the other for the probabilities of the several different pure states counted in the ensemble. The probabilities p_k in Eq. 26 are not necessarily the modulus squared of the eigenstate probability coefficients but represent the relative frequency with which the pure state $|\Psi_k\rangle$ in the ensemble occurs.

To clarify things beyond a mere abstraction, let us again take up the examples discussed above. As shown in the appendix (where we work out the right-hand side of Eq. 26), the density matrix of a particle always in spin-up eigenstate $|+\rangle$ is:

$$\rho(|+\rangle) = |+\rangle\langle+| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \quad \textit{Eq. 29}$$

One always has certitude (p=1) of finding the particle in the eigenstate $|+\rangle$. This is reflected by the single unit value on the upper diagonal element of the matrix. In a perfect analogy, for a particle in spin-down eigenstate, $|-\rangle$, the density matrix is:

$$\rho(|-\rangle) = |-\rangle\langle-| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \quad \textit{Eq. 30}$$

If one considers the two cases in a mixture of pure states, namely a statistical ensemble of eigenstates $\{(|\Psi\rangle_1 = |+\rangle, p = 0.5), (|\Psi\rangle_2 = |-\rangle, p = 0.5)\}$, for example, measuring a single particle from a stream of particles in mixed states and which can be found in one or the other eigenstate (because it is a mixture and *not* because of being the superposition of states!), then one sums up according to *Eq. 27* (or, again, see the appendix) and obtains:

$$\rho(|+\rangle, |-\rangle) = \tfrac{1}{2}|+\rangle\langle+| + \tfrac{1}{2}|-\rangle\langle-| = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}. \quad \textit{Eq. 31}$$

We are dealing with 2x2 matrixes because we consider the two-dimensional Hilbert space spanned by the two eigenstates $|+\rangle$ and $|-\rangle$. The diagonal elements of the matrix tell us that there is a 50% chance of measuring the spin-up or spin-down particle.

Now consider a single particle in a pure state – namely, in a spin-up and spin-down superposition. In contrast to the previous case, this is a pure state because it can be described by the single state vector $|\Psi\rangle = \frac{|+\rangle \pm |-\rangle}{\sqrt{2}}$. Then one gets:

$$\rho\left(\frac{|+\rangle \pm |-\rangle}{\sqrt{2}}\right) = |\Psi\rangle\langle\Psi| = \begin{pmatrix} \frac{1}{2} & \pm\frac{1}{2} \\ \pm\frac{1}{2} & \frac{1}{2} \end{pmatrix}. \quad \textit{Eq. 32}$$

We can extend this to entangled particles $|\Psi\rangle = \frac{|+\rangle_A|-\rangle_B \pm |-\rangle_A|+\rangle_B}{\sqrt{2}}$ or $|\Psi\rangle = \frac{|+\rangle_A|+\rangle_B \pm |-\rangle_A|-\rangle_B}{\sqrt{2}}$ (which, as you might recall, are the generalized particle version of the Type II or Type I entangled photons, respectively) and obtain:

$$\rho\left(\frac{|+\rangle_A|-\rangle_B \pm |-\rangle_A|+\rangle_B}{\sqrt{2}}\right) = |\Psi\rangle\langle\Psi| = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \pm\frac{1}{2} & 0 \\ 0 & \pm\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \textit{Eq. 33}$$

$$\rho\left(\frac{|+\rangle_A|+\rangle_B \pm |-\rangle_A|-\rangle_B}{\sqrt{2}}\right) = |\Psi\rangle\langle\Psi| = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \pm\frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \pm\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}; \quad \textit{Eq. 34}$$

respectively. This time we had a 4x4 matrix because of the four possible eigenstates $|+\rangle_A|+\rangle_B$, $|+\rangle_A|-\rangle_B$, $|-\rangle_A|+\rangle_B$ and $|+\rangle_A|-\rangle_B$.

Let us see what can be said by looking at the structure of these matrixes. First of all, again, the diagonal elements always tell us the probability of getting one or the other spin. Note that if one sums up all the diagonal elements of the density matrix, then one gets unity. Formally, this is described by the *'trace operator'*, Tr, on a matrix ρ, and is written as:

$$Tr(\rho) = 1, \quad \textit{Eq. 35}$$

Which, after all, is an obvious fact considering that the diagonal elements are the probabilities of all possible states.

But what do the off-diagonal elements mean? These off-diagonal elements appear every time one deals with entangled particles or superposition states or, in more general terms, when coherence and quantum correlation between subsystems is at work. There are none if a system is in a mixed decohered state but they pop up if some coherence is present or if the system (or sub-system) interacts with the environment. For example, for entangled particles, we might intuitively say that the value of the off-diagonal elements represents the 'strength' of the mutual entanglement. If they are all equal, one speaks of a *'maximally entangled'* system – something we already encountered when describing the Bell-states (see Eq. 15 and Eq.

16 in IV.1). This is the property of the density matrix which turns out to be very useful in describing – in a much more general manner than the wavefunction or state vector – all quantum systems, be they statistical ensembles of particles or one or many particles in superposition or entangled state.

All this has given us the basics necessary to distinguish between pure and mixed states from a formal point of view. What follows is a short review of some of the main algebraic properties of the density matrix. We won't prove them but they are worth mentioning, as if you intend to go beyond the reading of this introductory manual, you will find them mentioned everywhere in the specialized literature concerned with the foundational aspects of QP.

If you recall your high school math and remember that the product of matrixes (such as $\rho^2 = \rho \cdot \rho$) is arrived at by a row per column product, it can be shown that for each pure state, its squared density matrix is, again, the matrix itself, that is: $\rho^2 = \rho$. Therefore, as an obvious consequence of *Eq. 35*, the trace of the square of the density matrix is also what tells us whether a quantum system is in a pure state or a mixed state. Summarizing:

$$\rho^2 = \rho\,; \quad Tr(\rho^2) = 1 \rightarrow \text{pure state;}$$

$$\rho^2 \neq \rho\,; \quad Tr(\rho^2) < 1 \rightarrow \text{mixed state.}$$

Check this out by yourself with the above-given examples. $Tr(\rho^2)$ can also be considered a measure of the 'purity' of a state. If it attains a value close to one, it has a high degree of 'purity'.

The expectation value (that is, the average value one obtains by measuring an observable A many times) is:

$$\langle A \rangle_\rho = Tr(\rho A).$$

Density matrixes turn out to be very useful for expressing, in a compact fashion, a generalized form of the entropy of a quantum system. It can be shown that the Boltzmann-Gibbs entropy of Eq. 24 can be extended to quantum systems. This is the *'Von Neumann entropy'* (also *'entanglement entropy'*) and reads:

$$S = -Tr(\rho \cdot \log \rho).$$

If the entropy must be represented in bits of information, one uses the basis 2 logarithm. The matter of how to compute the logarithm of a matrix is beyond the scope of this treatise. (In the simple case of a diagonal matrix, it is simply the matrix with the log on each of its diagonal elements. For the passionate reader: Because, in linear algebra, there is a procedure that allows, by a change of vector basis, the diagonalizing of matrixes, each matrix has its correspondent logarithmic representation.) However, the point is that, as noted previously, for the density matrix algebra, it turns out that, again, a pure state always has zero entropy and a mixed state must always have some entropy:

$$S = -Tr(\rho \cdot \log \rho) = 0 \rightarrow \text{pure state;}$$

$$S = -Tr(\rho \cdot \log \rho) > 0 \rightarrow \text{mixed state.}$$

One can also show that, for Von Neuman entropy, the sum of the entropy of the subsystems is less than the entropy of the whole system. In formal language, quantum entropy is sub-additive. That is, given sub-systems A, B, C, … :

$$S(\rho_A + \rho_B + \rho_C \dots) \leq S(\rho_A) + S(\rho_B) + S(\rho_C) + \cdots$$

This is, again, a more general algebraic statement that the sum of the parts is less than the whole.

7. Measurement and information theory

This wired thing that a system in a zero-entropy pure state might well have subsystems that must be considered in a mixed state with non-zero entropy (against our classical intuition, which says that the number of microstates of the whole should be the product of all the microstates of its parts) forces us to the conclusion that, in QM, the whole is more than the sum of its parts. This, besides quantum entanglement, is frequently cited as an example of *'quantum holism'*. However, not every physicist would agree with this holistic interpretation, as, after all, what we see in our world is always the world in a collapsed and separable state.

This leads us to a delicate issue that has profound consequences. We would like to focus on it a bit longer because it will prepare the ground for the next chapter, about the black hole paradox and the holographic principle. The point is that every measurement on a pure state converts it into a mixed one, increases the entropy, and, translating this into an informational language, leads to a loss of information. A pure quantum state corresponds to a maximum information and maximum knowledge state that is allowed by QM. Zero-entropy means that the system's state is described by a single known state vector and that there is nothing more to be known – in contrast to a mixed state with more than zero-entropy, where the system must be described by more than one possible pure state for which we don’t know which of the alternatives are realized due to our ignorance of the system's configuration. In QM, a measurement is intrinsically a physical information erasure which removes the quantum interference between two possible quantum states. For example, a measurement on a particle in a superposition state sends the matrix of *Eq. 32* into the matrix of *Eq. 29* or *Eq. 30*. A physical measurement ‘cancels’ the off-diagonal values of the density matrix, as these are the coherence terms that measure the degree of quantum interference between two eigenstates. Therefore, a measurement introduces a fundamentally irreversible change. As we know from the second law of thermodynamics (recall the definition of irreversibility in chapter IV.5, Eq. 22), a transformation is irreversible whenever the entropy increases. There is no way to go back to the initial condition after a quantum system has been subjected to an information readout with an external measurement device.

However, as long as no measurement occurs (that is, as long as no other particles or physical objects interact with the quantum system in a pure state), there is no decoherence and, therefore, no information loss. The time-evolution of an isolated quantum system must, therefore, be determined by a unitary operator, which we encountered when dealing with QLG (and defined in Vol. I in the chapter on the state vector and Schrödinger equation). A unitary operator U applied on a state vector or, more in general, here on a density matrix at an initial time 0, $\rho(0)$ is a transformation that 'evolves' a quantum system from time t= 0 to time t and that preserves the total probability. Expressed in other terms, the density matrix is invariant under the action of a unitary operator as:

$$\rho(t) = U\rho(0)U^*, \quad \textit{Eq. 36}$$

with

$$U \cdot U^* = |U|^2 = I\,. \quad \textit{Eq. 37}$$

Where the star symbol, as usual, is the complex conjugation and I is the identity (matrix). From that follows the fact that the entropy does not change under a unitary transformation, as $S\big(\rho(t)\big) = S(U\rho(0)U^*)$. An evolution operator transforms a pure state into another pure state. This tells us that every evolution in time of an isolated quantum system must be unitary and preserve entropy information or, conversely, only an interaction or measurement that breaks the unitarity can lead to information loss. However, a quantum system that would evolve spontaneously from a pure state to a mixed state would be a physical impossibility that would violate the known laws of QM and of classical thermodynamics – namely, the second principle.

To summarize what we have learned in the last three chapters, let us restate the difference between a pure and a mixed quantum state with the following table.

Pure state	**Mixed state**
Single unit state vector $\|\Psi\rangle$ Examples: Eigenstates $\|\Psi\rangle = \|+\rangle$ or $\|\Psi\rangle = \|-\rangle$. Single particle superposition state: $\|\Psi\rangle = \frac{\|+\rangle \pm \|-\rangle}{\sqrt{2}}$ Entangled particles: $\|\Psi\rangle \frac{\|+\rangle_A\|-\rangle_B \pm \|-\rangle_A\|+\rangle_B}{\sqrt{2}}$ Bose-Einstein-Condensates	Statistical ensemble of pure states $\{ (\|\Psi_1\rangle, p_1), (\|\Psi_2\rangle, p_2), \ldots\}$ Examples: Ensemble of independent particles, noisy preparation of pure states, unpolarized light, etc.
$\rho^2 = \rho\,; \quad Tr(\rho^2) = 1$	$\rho^2 \neq \rho\,; \quad Tr(\rho^2) < 1$
$S = -Tr(\rho \cdot \log \rho) = 0$	$S = -Tr(\rho \cdot \log \rho) > 0$
Sum of the entropy of subsystems is larger than the entropy of the system	Sum of the entropy of subsystems is the entropy of the system
A measurement converts it into a mixed state	Can't be returned to a pure state

However, Stephen Hawking and Jacob Bekenstein, in a series of papers published between 1972 and 1975, showed that, according to our current understanding of Einstein's relativity, BHs seem to be able to violate unitarity: A quantum system falling in the gravitational field of a BH may spontaneously transform from a pure state into a mixed one and, therefore, violate the physical laws as we currently understand them. Again, this is a typical anomaly, as the history of science has frequently encountered, and has led to an apparently endless controversy through today. We will review it in the next chapter.

8. Black hole thermodynamics

You have certainly heard about this mysterious object which astrophysicists call a *'black hole'* (BH). When people think of physicists, most of the time they connect them to nuclear weapons or BHs. We prefer to focus on the latter. Though this is not a book on Einstein's relativity, let us try to summarize it, nevertheless, with a quick description of what BHs are about.

As you might already know, in astrophysics, BHs are those kinds of peculiar objects which are the final remnants of a star which has run out of its nuclear fuel. Once a star, after hundreds of millions or billions of years, depending on its mass, has exhausted all its internal energy, only a few possibilities might prevent its gravitational collapse onto itself. If the star's mass is smaller than 1.4 $M_\odot$ (1 $M_\odot$ = one solar mass), the so-called *'Chandrasekar limit'*, it will shrink to the type of white dwarf star we already discussed in the chapter on fermions and bosons in Vol. I. The Pauli exclusion principle prevents further collapse because the degeneracy pressure still holds up the entire mass of the almost-dead star. Meanwhile, for stars between 1.4 $M_\odot$ and 2.8 $M_\odot$, the gravitational field becomes strong enough to overcome the degeneracy pressure but not that of the strong nuclear forces that act among protons and neutrons. Eventually, what remains of the star after a supernovae explosion will collapse to an object which is dense enough to literally squeeze together protons and electrons to form a unique nuclear exotic state of matter made almost exclusively out of neutrons, wherefrom it got its name: *'neutron star'*. However, the final destiny of stars that, after eons of radiation and matter emission and eventual novae or supernovae explosions, still contain more than about $3M_\odot$ is unavoidable: These are doomed to an unstoppable catastrophic collapse into a BH.

These extreme objects owe their name to the simple fact that their gravitational field curves space-time so strongly that even light, if it comes sufficiently near to it, is no longer able to escape and circulates for an infinitely long time in its orbit. Or, if it comes even nearer, it eventually falls into the center of the BH and will never be able to climb back up the gravitational force field. The latter limiting region is called the *'event horizon'* (EH) because any event happening inside it will never be detected from the outside and there is no possibility of communication between it and the outer universe. Anything or anyone falling into the BH, say astronauts on a spaceship, and traversing the EH is lost forever. They will have no chance to ever get out from the BH because, to do that, their spaceship would have to accelerate to an escape velocity greater than the speed of light – that is, they would need an infinite energy, something which, as we know well, is strictly forbidden by relativity. The EH also defines the size of the BH and, in general, is determined by the mass, charge, and angular momentum of it. For the simplest case, that with no charge and angular momentum, which we assume to be a good approximation of most existing BHs, the EH depends only on the mass and acquires a spherical shape. In this latter case, the radius R_S of such an EH, also called the *'Schwarzschild radius'*, can be easily calculated by simple principles and turns out to be (in units of meters):

$$R_S = \frac{2GM}{c^2}, \qquad \textit{Eq. 38}$$

With $G = 6.674 \times 10^{-11}\frac{N\cdot m^2}{kg^2}$ Newton's universal constant of gravitation, M the mass of the BH (in kg), and $c = 3 \times 10^8\,\frac{m}{s}$ the usual speed of light. The reader is invited to play with this little formula (using a scientific calculator) to see that, for instance, a BH of 3 $M_\odot$ would have an EH of about 8.8 km radius. Or, if we would compress the Earth into a BH, it would have a radius of 9 mm!

In the vicinity of the EH, light is deviated, bent, or scattered. In general, every gravitational field deviates the light path of a photon. If the body is a spherical one, it can also act like an optical lens. This effect, which is not only present for BH but is also characteristic of any material body with a sufficiently strong gravitational field, is called *'gravitational lensing'* and is amply used by astronomers to determine the matter distribution in galactic clusters. Obviously, it becomes particularly strong in the case of a BH, as illustrated in Fig. 64.

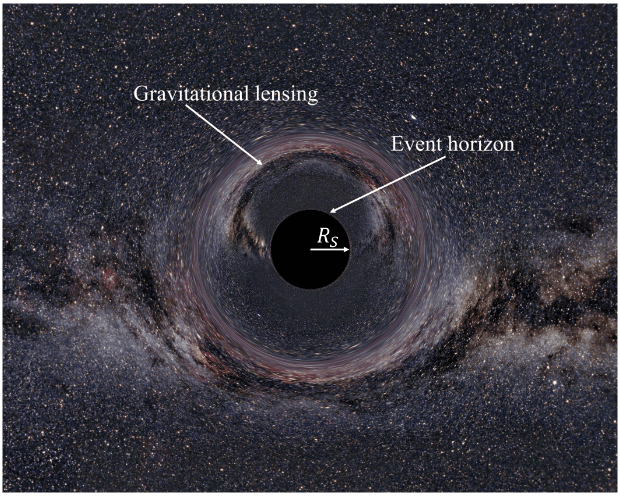


Fig. 64 Simulation of a BH gravitational lensing against a background of stars.

In principle, a BH is a relatively simple object. All the matter and energy falling into it have the same fate: Anything with a structure is compressed and finally destroyed without any possibility of coming back. A BH is, therefore, expected to be made of matter which is substantially in the same state everywhere and can be described by few physical parameters, such as its mass, electrical charge, and angular momentum. According to this hypothesis, summarized by the so-called *'no-hair theorem'*, a BH was supposed to be simply a totally black, inactive, smooth object without any other properties, and even having zero entropy.

However, because the compression of such huge amounts of matter into so tiny a volume leads to extreme physical conditions and extreme gravitational force fields, this theoretical abstract boundary around a BH is also a boundary between the laws of physics as we know them – namely, QP and GR – and a new physics of QG which is still something physicists have to find, as we discussed in chapter III.3.

Therefore, BHs are interesting astrophysical objects insofar as they represent a theoretical testbed for our known physical theories. Nearby, a BH space-time is so curved that certainly no classical Newtonian physics holds. Also, Einstein's GR works only until a certain limit. In addition, QM and QFT are certainly no longer a completely valid theory under these conditions. As we will show, it is nowadays clear that neither GR nor QM accurately describe the physics near the EH and even less inside a BH, as their direct application leads to paradoxes and contradictions such that only an as-yet-unknown theory of QG will eventually be able to do that.

However, this doesn't mean that some interesting things can't already be said about quantum phenomena in curved space-time. Moreover, taking a look at the contradictions and paradoxes that arise in trying to apply QP to BH is certainly an interesting exercise that clarifies where modern science stands and where truth might be hiding.

To see where this leads us, let us first set the stage by describing a quite counterintuitive quantum phenomenon that almost certainly characterizes BHs – namely, the *'Hawking radiation'*. To see why BHs are perhaps not so black after all, consider an observer who is accelerated relative to another non-accelerating observer. Physicists call any frame of reference which is not accelerating (i.e., no force is acting on it) an *'inertial frame of reference'*, while any accelerating one (i.e., a force is acting, such as a gravitational field on the infalling observer) is a *'non-inertial frame of reference'* (or just an inertial or non-inertial detector or 'observer', respectively). Between 1973 and 1975, Stephen Fulling, Paul Davies, and William George Unruh showed, in a series of papers, that the zero-point energy of the vacuum is not an absolute concept but, rather, depends on one observer to another. Modern QFT shows that the ground state of the vacuum, which is the lowest possible quantum state, must be different for a non-inertial reference than that of the inertial one. An observer who is accelerating 'sees' the same vacuum with a little bit higher energy ground state as an observer who does not accelerate. Unruh showed that it must appear in the form of a black body radiation. The curious result is that the accelerating observer will be surrounded by a gas of particles. This is the 'Fulling–Davies–Unruh' effect, or, in short, just the *'Unruh effect'*, which predicts that a non-inertial detector will measure a black body radiation with a temperature proportional to the acceleration, while an inertial detector will measure none. The 'observer' may also be a thermometer: If a non-accelerated thermometer measures zero Kelvin, the accelerated one will always measure a temperature slightly higher than absolute zero.

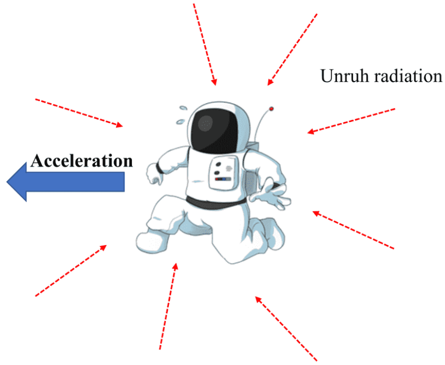


Fig. 65 The Unruh effect illustrated from the standpoint of a non-inertial observer.

To date, this remains a speculative theory. Why this is the case becomes clear when we try to estimate the magnitude of this effect. The black body temperature (in Kelvin) of the *'Unruh radiation'* is given by the simple formula:

$$T_{Unruh} = \frac{\hbar a}{2\pi c k_B}, \qquad \textit{Eq. 39}$$

With a the acceleration (in m/s^2) and, as usual, $\hbar$ the reduced Planck constant $\frac{h}{2\pi}$, c the speed of light, and k_B the Boltzmann constant. (See the list in Appendix A VI.) For the acceleration of 1g, that is, the acceleration at the Earth's surface ($a = g = 9.81\ m/s^2$), one gets a temperature shift of about 4×10^{-20}K, which is such a tiny difference in temperature that it is beyond the sensibility of any thermometer. To obtain an Unruh temperature of 1K, an acceleration of about $2.5 \times 10^{20} m/s^2$ (about a billion of billion times the Earth's gravity acceleration) is necessary; this is such a huge acceleration that no conceivable laboratory experiment with a physical macroscopic object like a thermometer or detector could recreate it and survive the impact of the acceleration force without being smashed and torn into pieces. Though the experimental proof is beyond the reach of present technology and there is no universal consensus (some physicists doubt its existence and some aspects remain a matter of debate and controversy), the Unruh effect is, however, taken as very plausible. If it exists, the Unruh effect remains a completely negligible effect for our everyday experience, as long as we are not dealing with objects subjected to extreme physical conditions.

However, what happens in an environment of extreme physics like that near the EH of a BH? The gravitational force of attraction at a BH boundary can cause such accelerations that the Unruh radiation might become relevant. In fact, Einstein's GR, which is the best-known theory we have for describing BHs, is based on the *'equivalence principle'*. Loosely speaking, the equivalence principle of GR states that all inertial forces are equivalent in the sense that a uniformly accelerating reference frame is indistinguishable from a non-accelerating reference frame in a gravitational field. Another common way to illustrate the same principle is to consider *'Einstein's elevator'* of Fig. 66.

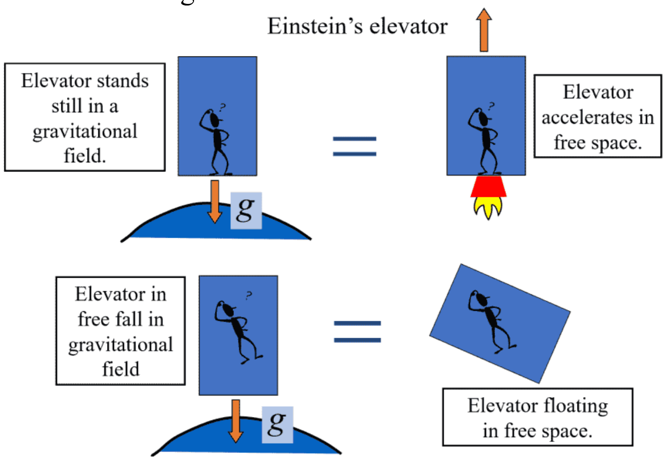


Fig. 66 Einstein's thought experiment to illustrate the equivalence principle of GR.

Consider two observers in weightlessness: one inside a stationary elevator and another inside the same elevator far from any gravity source but accelerated by an external force (say, a rocket). That means the gravitational acceleration of a body is equivalent to a non-inertial frame reference without the gravity field – or, equivalently, one inside an elevator falling in a gravitational field and the other inside the same elevator far from any gravity source floating freely without an external force changing its state of motion. From the point of view of an inside observer, there is no way to distinguish the two cases, not even in principle. These two situations can be considered physically exactly equivalent. The validity of the equivalence principle is beyond question because GR is a theory that has been confirmed repeatedly by experimental evidence throughout the last century.

Now, this also suggests that an astronaut falling into a BH should feel nothing and will even traverse the EH without noticing anything (an eventuality called the *'no-drama scenario'*). Only to an outside observer does the EH represent a limit without return. Strictly speaking, the absence of drama is not entirely correct because, especially for BHs with smaller masses (smaller, not larger!), colossal tidal forces will tear into pieces any object with a spatial extension.

Tidal forces arise due to the difference in the gravitational force between two nearby points in space, such as between the feet and the head of an astronaut falling into it. However, gravitational tidal forces are differential *'apparent forces',* a completely different kind of force which acts only on bodies with a size. They to do not contradict the equivalence principle of GR. In fact, simply take a point-sized particle (that is, without a size and extension) and any tidal force vanishes.

Therefore, theoretically at least, this is no issue: There is no physical difference between a particle in free fall, even near the EH of a BH, and a particle far enough from any gravitational force in weightlessness. The same fits for the equivalence between an observer who is accelerating with a spaceship near an EH of a BH, in order to avoid falling into it, and the same observer accelerating with a spaceship but far from the BH.

This equivalence must also hold for the Unruh radiation. The astronaut near the EH igniting a powerful rocket which allows the spaceship to remain at rest just above the EH can, nevertheless, be considered as accelerating. An external observer will see the Unruh radiation surrounding the spaceship in a thermal bath.

However, because the EH itself can be considered an accelerating reference frame, the natural question is: Will an astronaut at a safe distance from the BH see the BH radiating the same Unruh radiation from its EH? If the equivalence principle holds (and it holds in all the physical situations we know of, as otherwise GR would not have been such a successful theory), then one must answer affirmatively: BHs radiate energy and cannot be as black as was previously assumed.

The achievement of Stephen Hawking, the world-famous British physicist and well known to the audience for his popular science books, was to show how one obtains the very same result without resorting to the Unruh effect but, rather, by using a completely different approach. Moreover, Hawking could show that this implies that BH must lose mass with the passing of time and that there must be something intrinsically incomplete with our conceptions of relativity tied to quantum phenomena.

Fig. 67Stephen Hawking (1942-2018)

Hawking used QFT on curved space-time at the EH. He tried to understand how the quantum 'fluctuations' of the vacuum close to the EH of a BH behave. What he found was that the virtual pairs of particle and anti-particle creation filling the vacuum (see the chapter on zero-point energy, virtual particles, and the Casimir effect in Vol. I) near the EH may behave in a dramatically different way than they would in space far from any too-strong gravitational field. As we know, virtual particles can be visualized as energy quanta that pop in and out of existence very quickly. They are not considered 'real' but only 'virtual' because they are allowed to exist for only an incredibly short period of time determined by the time-energy uncertainty principle and must then literally 'disappear'. Otherwise, they would violate the energy conservation principle. One can imagine this as a process in which a foam of entangled virtual pairs of particles appears and begins annihilating each other shortly thereafter (which, in the case of material particles, would be a particle and anti-particle pair), leaving no trace in the empty 'real' space. The point is, however, that if an entangled virtual particle pair appears sufficiently near the EH, it may eventually occur that one of the particles traverses it and falls into the BH without any chance to return and leaves the other entangled particle behind in a still-safe zone that it can, in principle, escape but with its counterpart no longer being available with which it can annihilate and 'disappear'.

Let us restrict ourselves to the simple case of photons. (A similar process occurs for material particles but is much less effective.) This means that the space-time near a BH is so strongly curved that it can, so to speak, 'tear apart' the entangled virtual photon pairs coming into existence at the EH surface and transform them into real photons. One would be lost in the BH forever, out of the sight of any outside observer, but the other photon can escape.

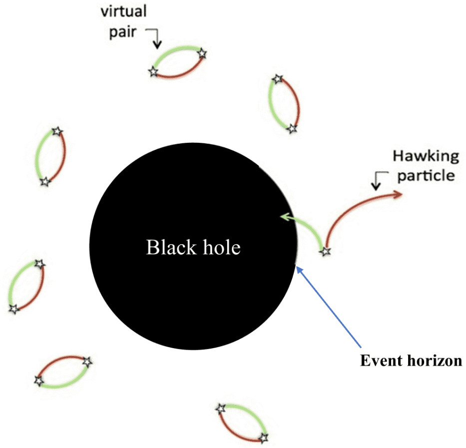


Fig. 68 Hawking radiation from a black hole.

The known laws of QP also allow the same process to take place near the edge of the internal part of the EH: One of the virtual entangled photons may quantum tunnel through the EH, 're-appear' on the other side, and escape the BH. An outside observer would, indeed, see the BH radiate a weak but real gas of photons – the so-called *'Hawking radiation'*.

Detailed calculations have shown that it is a black body radiation whose temperature is given by exactly the same formula as the Unruh radiation. Hawking adopted another quantum theoretical path and interpretation but finally confirmed that BHs must radiate energy. The Hawking radiation and the Unruh radiation are just two sides of the same coin.

However, this was not the whole story. The question is: How can a virtual particle become real and suddenly come into existence when, previously, nothing was there? Where are this 'new' mass and energy supposed to come from? The simple answer that Hawking gave is that it is the BH itself that loses its mass. As we know well, due to Einstein's mass-energy equivalence principle, mass can also be converted into energy and vice versa. If we assume that, for every particle escaping the EH, the BH lowers its mass correspondingly (that is, Hawking radiation causes a decrease in the BH's mass), then no issue with the energy conservation arises. In other words, due to the Hawking (or Unruh) radiation, BHs are destined to 'evaporate'.

In most cases, the BH evaporation is an extremely slow process. One can show that the black body temperature associated with the Hawking radiation is inversely proportional to a mass M as:

$$T_{Hawking} = \frac{\hbar c^3}{8\pi\, k_B G M} = \frac{6.16\times 10^{-8}}{M_{\odot}}\ K\,, \qquad \textit{Eq. 40}$$

with the right-hand side of the equation expressing the temperature as a direct proportionality factor for solar masses. (For a simple derivation, see Appendix A IV.) Because, for a one solar mass BH, the temperature is as low as 6.16×10^{-8} K (about 61 nanokelvin), it becomes clear that the radiation is quite weak. This temperature is several orders of magnitude smaller than the cosmic background radiation which is, at the present age of the universe, about 2.7 K. (Recall the end of the chapter on the blackbody radiation in Vol. I.) More radiation will fall into the BH than it emits, and the net result is that it will increase in mass despite its loss due to Hawking radiation. Only in a very distant future, once the universe has expanded enough to dilute the cosmic background to temperatures lower than that, will the BH begin to decrease its mass. Then it will still need other eons to do that. It can be shown that a BH evaporation time scale is proportional to the cube of its solar masses as:

$$t_{evap} = 2.11\, M_{\odot}{}^{3}\ \times 10^{67}\ y\,.$$

It will take more than 10^{67} years for a solar mass BH to evaporate! These are absolutely insane time scales – much greater than the age of the universe.

It is, therefore, fair to say that a BH is nevertheless an eternal object. However, the evaporation time of a BH with a mass of the order of 10^{11}kg (that is, about the mass of a tiny asteroid) would be about the time corresponding to the age of the universe (ca. 13.7×10^9 years). This led astrophysicists to speculate that, if such small BHs were formed during the first instants of the Big Bang, we should nowadays detect the footprint-signal of so-called *'primordial black holes'* in its last quantum evaporation phase. So far, no such evidence has been found. This might sound like a falsification of Hawking's theory. However, a possible alternative explanation for the lack of such evidence is that, likely, these BHs did not form during the Big Bang in the first place. However, if, for example, someone smashes two protons against each other with sufficient energy, it is conceivable that they will form a microscopic BH which will evaporate almost instantaneously. This is also what powerful accelerators probably do continuously. There has been some concern, and a wide portrayal by the media, that the LHC could produce a BH that might swallow the entire Earth in a catastrophic doomsday scenario. However, this can't be the case for at least two good reasons. First, the evaporation time of a two-proton mass BH is so small (about $10^{-100}s$!) that it has no time to encounter any particle and do anything; it can't go anywhere. Secondly, much less theoretical but very empiric evidence shows that, above our heads, the atmosphere is permanently bombarded by cosmic rays containing particles that are hundreds of millions of times more energetic than what the LHC can produce. And yet, we are still here! So, there is nothing to fear.

9. The black hole information paradox and the holographic principle

Hawking radiation is interesting in many other respects. One of the things Hawking soon realized is that there must be something fundamentally flawed with our understanding of quantum and/or relativistic phenomena.

Let us focus again on the physical process occurring at the boundary of the EH. According to Hawking's model, here we have virtual entangled particles which are converted near the EH of a BH into real non-entangled particles. Recall what we concluded in section IV.7 about unitarity. A pure state can never spontaneously become a mixed state, as this would violate the time evolution unitarity. If there is no one or nothing that measures and interacts with a pair of entangled particles, QM tells us that these must be described by a single wavefunction which can't collapse entirely on its own to a mixed state of two particles described by the density matrix. Unfortunately, however, this is precisely what, according to our understanding of QP and GR, the Hawking radiation model seems to suggest. At the EH, the two entangled virtual particles in a pure state are suddenly converted into a mixed state of two real particles (sort of like the objective collapse theory we described in II.4), one falling into the BH and cut out from the outside universe forever, and the other contributing to the gas of photons making up the Hawking radiation. Formally, one could compare this with a spontaneous collapse of the entangled photons represented by the density matrix of Eq. 33 to that of the mixed state pair of the particle density matrix of Eq. 31. However, in chapter IV.7, we emphasized that precisely this sort of transformation can't exist without an external act of measurement, as otherwise we would violate the second law of thermodynamics. We showed that the density matrix must be invariant under the action of the state evolution operation. Otherwise, such a 'spontaneous breakup' would violate the unitarity that is expressed in Eq. 36 and Eq. 37.

So, something must be badly wrong here – something which evidently escapes our full understanding.

One might argue that the intense gravitational force acting on particles near the EH of a BH could be considered, in all respects, a very strong interaction and that, therefore, we should not be surprised that, due to this ‘natural measurement’ process, the Hawking particles in a pure state are transformed into a mixed one. However, this also implies a direct violation of the equivalence principle in GR. We have already pointed out how any local observer in free fall in a gravitational field would feel nothing unusual in traversing the EH. From the point of view of the particles falling freely in the BH, there is no such force acting on it in the first place.

Therefore, one might be tempted to believe that GR is not a complete theory and must be reconsidered. However, because the equivalence principle is one of the main pillars of GR, and because no known physical process has so far contradicted it, physicists are not willing to give it up so lightheartedly. On the other hand, it might be our present understanding of unitarity in regard to the evolution of pure and mixed states described by the density matrix formalism in QP that must be amended. However, unitarity is also the main pillar of QFT and the SM of particle physics, which, as we have pointed out in a previous chapter, is one of the most successful theories ever. Perhaps both theories need to be updated to a future theory of QG that solves this apparent dilemma. However, this has not yet been discovered.

Another way to illustrate the same problem is through the information-theoretical standpoint. We saw that the entropy of a pure state is zero – that is, a pure state is described by a maximal amount of quantum information. Meanwhile, the entropy of a mixed state must always be greater than zero – that is, there is some incertitude about its configuration, the number of possible microstates. Because the Hawking radiation process seems to convert a quantum pure state into a mixed state, this also implies that the total entropy (outside and inside the EH) has been increased and that some information content must have been lost forever in the universe. We might say that BHs destroy information. Where has this information gone?

Note that this is a very different scenario from that which follows from the second law of thermodynamics. While it is true that the second law tells us that the entropy in an isolated system must always increase, in principle there is no physical reason or law that does not allow an open system to locally decrease it again to the expenditure of the total entropy of the universe and bring it back to the initial state. The typical example could be a hot reservoir with high entropy decreasing its entropy by transferring its heat to another colder reservoir.

By the way, life continuously resorts to this. Our living bodies and all our cells can be considered thermodynamic heat reservoirs that avoid entropy increase and even decrease their entropy by exchanging heat with the environment, increasing, however, the total entropy of the surroundings. What we call 'death' is the physical failure of an organism to control its local entropy balance.

However, the laws of physics are inherently time-symmetric and there is no reason why the same phenomenon can't coincidentally go the other way around, just like, with a movie, one can observe a scene again by playing back the film. Why can't the heat be transferred back and exactly reconstruct the initial state (or a dead body resurrect)? This irreversibility is what, in essence, the second law of thermodynamics means. It is the so-called problem of the *'thermodynamic arrow of time'* which asks why, despite the fact that nothing in the laws of physics is preventing physical processes from going backward in time, we nevertheless live in a universe where we never observe such a time reversal of physical phenomena. The solution to that dilemma, and the one which physicists usually put forward, is that the probability that the incredibly huge number of molecules (recall Avogadro's number!) dominated by a thermal molecular chaos may eventually, just by coincidence, follow the same path back, in the opposite direction, is so incredibly low that we will nowhere and never observe such an eventuality in the entire history of the universe. That's why we will never see a cold body in contact with a hot one becoming even colder or (to make a less creepy example) see the pieces of a broken glass reverse the shattering process and come back together. However, in principle, it is not a physical impossibility. In the information-entropy analogy, this means that, at least in principle, one can recover all the previously lost information by tracing back the position and movement of every single constituent (atoms or molecules) of the thermodynamic body and reconstruct its exact previous whereabouts – a possibility that resonates well with the good old Laplacian determinism.

However, in the case of quantum processes in which a violation of unitarity with a spontaneous projection of a quantum pure state into a mixed one seems to take place, there can't be any such back-tracing as in classical mechanics. We know that quantum probabilities are very different from classical probabilities. The latter are due to our ignorance, while the former are intrinsic. No matter how much care and precision we use to re-establish the original boundary conditions, we will never be able to 're-entangle' two particles and be sure that they will show up with the same spin values as they did in the previous measurement. The 50% chance for the two possible outcomes remains and is unavoidable. This also implies that we will never be able to play back the film as one can do, at least in principle, in CP.

Therefore, a BH that does not preserve quantum unitarity: Transforming the entangled virtual particles at the edge of the EH from pure states into a couple of real particles in a mixed state does not preserve information. It increases entropy, not due to the second law but because it literally 'eats up' information, making it disappear from the universe.

All this, however, is contrary to our present understanding of QP and GR, according to which information never gets truly lost. At least in principle, it could be recovered. A paradox emerges, which is nowadays widely known as the *'BH information paradox'*.

The popular media illustrates the BH information paradox by resorting to the example of throwing books into a BH. Their information content will be destroyed forever. However, this information loss should not be interpreted as the destruction of the information medium, say, by burning the book. As Laplace liked to believe, because the known physical laws are time-symmetric, if one measures the exact whereabouts of each piece of burnt paper and considers the position as the dynamical boundary conditions of each molecule of the smoke, the flame, and any remnant emerging from the book's incineration, it is, in principle, possible to reconstruct backward in time (say, by a powerful and detailed computer simulation) the exact molecular structure of the book and its ink molecules with which the pages were printed and, consequently, its information content. However, this is precisely what the BH information loss and paradox is *not* about. The information contained in a book that traverses the EH of a BH is destroyed and lost forever because of this quantum-mechanical unitary violation; there is no way back that allows us to recover it. In a certain sense, speaking of information loss is misleading and, after all, does not tell us much. To use the play-back analogy of a movie, we can say that what is lost is the possibility of playing back the process which, for some reason that remains unclear, transforms from a reversible into an irreversible one. If, in CP, one is allowed, at least in principle, to play-back in time the events, and if this remains true in QP as long as it does not deal with strongly curved space-time, in the presence of a BH gravitational field where the laws of GR take over, the loss of unitarity sets in, making the process irreversible – that is, we are prevented from playing-back the film, not even in principle.

A first attempt to come up with this information paradox, aimed at saving physics from quantum information loss and unitarity violation, came in 1993 from the American Stanford University string theorist Leonard Susskind, who posited a *'BH complementary principle'*. According to Susskind, the contradiction disappears if we suppose that information is both inside and outside the BH.

From the quantum-microscopical perspective, the couple of virtual particles that become real, with one falling into and through the EH and with the other escaping the BH, will remain in an entangled state, maintaining a pure quantum state – that is, with no information loss or unitarity violation. The particles entering the BH are entangled with the EH. The idea that particles remain entangled with each other even through the EH does not lead to any contradiction with the prohibition that particles inside and outside a BH are not allowed to be in contact, as we know that there is no way, not even in principle, to use quantum entanglement as a means of transmitting information.

From the point of view of observers, this translates into the conjecture that that the information is accumulated and encoded first near the EH and, later, radiated away by the outgoing Hawking radiation. Because there is no possibility that an observer who has crossed the EH can communicate with the outside universe, there is no real paradox. For an external observer, the information falls into the BH and is first uniformly distributed and scrambled over a thin membrane slightly above the EH, which will later be radiated away into the external universe with no information loss or unitary violation. The observer's attempt to determine whether the glowing membrane is real by falling freely through this slightly stretched horizon will see it disappear. For the internal observer, no detectable change in the infalling matter is observed because, according to the equivalence principle of GR, the EH is simply an abstract theoretical boundary and not a physical object. An inside observer detects the information-entropy entering somewhere at the EH and does not notice anything special. However, because there is no way to report this to the outside, there is, in this sense, a complementarity between observations made by infalling observers who cross the EH and those made by distant observers.

Another aspect that suggests how our present understanding of the fundamental laws of physics break down in BH thermodynamics emerges when we explicitly calculate the entropy. Jacob Bekenstein, an Israeli-

American physicist well known for his contribution to BH thermodynamics, showed, among other things, that that the entropy of a BH of mass M is (the skilled reader can find a quick derivation in Appendix A IV):

$$S_{BH} = \frac{k_B c^3}{4\,\hbar G} A \propto \frac{1}{4} A\,, \qquad Eq.\ 41$$

Where A is the area of the BH (by definition, the area of its EH; of course, the suffix 'BH' may stand for 'black hole', though in the literature it also stands for *'Bekenstein-Hawking entropy'*). This is also the maximum entropy content that a sphere of surface A can contain.

The attentive reader might already glimpse the problem with this expression of entropy. In fact, nowhere in CP does the entropy of a body emerge as a function of its surface area. Entropy is a thermodynamic state function that depends on the mass and, therefore, the volume of the object it describes. If we take Boltzmann's statistical interpretation of entropy as a measure of the (logarithmic) number of possible microstates (that is, Eq. 23), it is quite hard to imagine it as being dependent only on the surface area. All the particles, with all their possible configurations, on the surface (such as inside the volume of a body) are supposed to contribute to the overall number of degrees of freedom. This makes the Bekenstein-Hawking entropy weird and is, again, another sign that something fundamental is escaping our understanding. Since Bekenstein published his result in 1973, generations of physicists have tried to come up with a reasonable interpretation.

A possible solution came in 1995 from the Dutch physicist and Nobel Laureate Gerard 't Hooft. He conjectured that the strange area-dependence of the entropy of a BH might indicate that we live in a *'holographic universe'* – that is, a universe in which the three spatial dimensions are a sort of illusion, being only the projection of a two-dimensional reality. Susskind refined this idea, giving it a precise mathematical background by means of a string-theoretical interpretation.

Most of us are familiar with popular holographic pictures. These are reproduced by a two-dimensional photographic film on which a couple of beams of coherent laser light are reflected and, by interferometric means, reconstruct the three-dimensional image. Holograms contain all the information about the object whose image must be projected. In general, this fact shows how all the information about an N-dimensional body can be stored in a N-1 dimensional object that functions as a total information carrier. Taking this as an analogy, Hooft conceived of our 4D universe (three spatial dimensions + one temporal dimension) as the projection of a 3D universe (two spatial dimensions + one temporal dimension, see Fig. 70). According to the *'holographic principle'*, all that is happening in a volume of space in time is, in reality, encoded in a 2+1 dimensional boundary which is, itself, physically equivalent to an EH (a *'light-like boundary'*).

Unfortunately, in the frame of this holographic principle, the 2+1 dimensional physical reality does not at all contain the force of gravity! Interestingly, however, a couple of years later, the Argentine physicist Juan Maldacena was able to show how this is in line with ST. Under particular conditions, gravity emerges as a lower-dimensional description of a higher-dimensional theory.

Fig. 69 Gerard 't Hooft, Leonard Susskind and Juan Maldacena.

How this could be is a too-long and too-complicated story that would need a separate treatise to be elucidated in the context of ST and QFT. It may, however, be said that a *'duality'* exists between an ST 'living' in a so-called *'anti-de Sitter space'* (AdS) and the conformal field theories (CFT), which describe the known QFT by a scale invariance approach. (The physics of the theory remain invariant at all length scales.) An AdS represents a non-Euclidean universe in line with GR but with a negatively curved space-time – that is, a decelerating expansion. (Our universe possesses a positive curvature due to the increasing expanding rate induced by dark energy.)

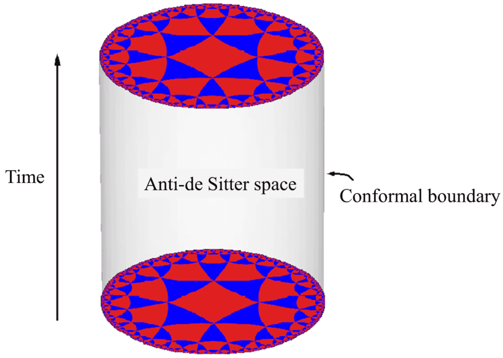


Fig. 70 The 3D AdS space-time looks like a solid cylinder.

Non-Euclidian means that, on the extreme cosmological scales, our classical Euclidian notions of distance no longer hold and must be modified to a hyperbolic space geometry where, for example, triangles and squares look like distorted and stretched objects as in Fig. 70 and with the cylindrical outer conformal boundary infinitely far from any point in the interior. Each

2D disk represents the universe at a specific instant in time that flows along the vertical time-axis forming a cylindrical structure.

Maldacena could prove that the boundary of the AdS without gravity can be considered equivalent to a space-time in a CFT with gravity. Moreover, a duality exists which transforms the stronger (weaker) force of one theory into the weaker (stronger) force of the other theory and whereby a string on the boundary of the former theory is a particle in the latter one. Every physical event can, thus, be described in a universe described by ST in an AdS space-time or, equivalently, a QFT on the boundary of that space-time. They both describe the same physics. This is also called the *'AdS/CFT-correspondence'* while the duality is termed *'holographic equivalence'* or *'holographic duality'*.

This correspondence ended up solving the entropy-information paradox conceptually and quantitatively. It can be explained illustratively by using the ultimately small space-time units that QM and GR allow for – namely, the *'Planck-units'*, which we already considered in chapter III.3. The Planck scale units are solely determined by the fundamental physical constants and represent the space-time scales where quantum fluctuations are intense enough to form micro-BH for an extremely short period of time. It turns out (see a not-too-rigorous derivation in Appendix A V) that these units are the following. The Planck length

$$l_P = \sqrt{G\hbar/c^3} = 1.6 \times 10^{-35} m \,, \qquad \textit{Eq. 42}$$

then Planck time

$$t_P = \frac{l_P}{c} = \sqrt{G\hbar/c^5} = 0.54 \times 10^{-43} s \,,$$

and the Planck mass

$$M_P = \sqrt{\hbar c/4G} = 1.1 \times 10^{-8} kg \,.$$

There are also other Planck units, but let us fix the ideas on these. If you inspect the order of magnitude of these values, you will quickly realize that we are talking about something beyond any human comprehension. For example, if we consider the size of an electron as being 1000 times smaller than that of a proton (say, 10^{-18}m), then the Planck length is still 17 orders of magnitude smaller (i.e., hundreds of millions of billions of times smaller). The LHC – to date, the most powerful particle accelerator – can probe distances as small as 10^{-19}m and upgrades or future accelerators might reach 10^{-20}m but we can set aside any hope of looking down at 10^{-35}m sizes. Unless there are some unexpected technological breakthroughs in building new particle accelerators (and nobody has a clue so far, but the history of science can be surprising in this regard), the possibility of investigating such horrendously small scales will be barred from us for generations. The Planck time scale is even further from anything we know. The most precise atomic clocks to date are able to measure time intervals of about 10^{-16}s (about a ten-thousandth of a trillionth of a second) but a Planck time interval is still 10^{27} times smaller! The Planck mass is the mass of the virtual BH appearing for a time interval as short as the Planck time and weighs about a hundredth of a milligram. The latter seems to be a more manageable quantity but if we consider that it is concentrated within a volume of a sphere with a diameter of the Planck length, then its density is about (check it as exercise) $5 \times 10^{87} \frac{kg}{mm^3}$. With this density, the entire matter of the known universe could easily be packed into a volume much smaller than a cubic millimeter!

All this should make clear that we are dealing with scales and their associated physical phenomena that are far beyond our actual technological and cognitive potentials. However, there are good reasons to believe that these are also the ultimate constituents, the fundamental 'bricks of reality', so to speak. ST, as well as other speculative theories, posit the Planck length and time as the single and smallest fundamental cell that makes up space-time. In this view, all physical reality is built up like a giant puzzle, the indivisible unit of which is a volume (or a surface) element with Planck length sides and where every 'time-tick' is as short as the Planck time. In other words, it is assumed that we live in a universe in which elementary space cells are Planck volumes (or surfaces) that are played out, one after another, like the separate frames of a motion picture film, each with Planck time duration.

Within this framework, the AdS/CFT-correspondence reinterprets the area dependence of the BH entropy. Note how, if one rewrites Eq. 41 in terms of Eq. 42, the entropy of a BH becomes (see also end of Appendix A IV):

$$S_{BH} = \frac{k_B A}{4\, l_p^2} \propto \frac{A}{4\, l_p^2}.$$

This tells us that the entropy of a BH is directly proportional to one-fourth the number of 'Planck tiles', having a Planck area of l_p^2 and covering the EH of area A, suggesting that the EH possesses a coarse graining of microphysical degrees of freedom. From Boltzmann's entropy we know that any entropy S is a measure of the number of microstates W a thermodynamic system can attain as $S = k_B \log W$; therefore, we can interpret the number of degrees of freedom of a BH as one-fourth of the number of Planck tiles covering its EH. One can, however, also take the information-theoretical perspective that we elucidated in the section on quantum information theory. In fact, this can be interpreted equivalently by using Shannon's entropy for the equally probable events of Eq. 25 as S = $log_2 W$ – or, in other words, a BH contains an amount of information S expressed in bits and the number of these bits is given again by one-fourth of the number of the Planck tiles covering its EH. Fig. 71 illustrates how four Planck tiles on the surface of the EH of a BH represent one unit of entropy – that is, one bit of information.

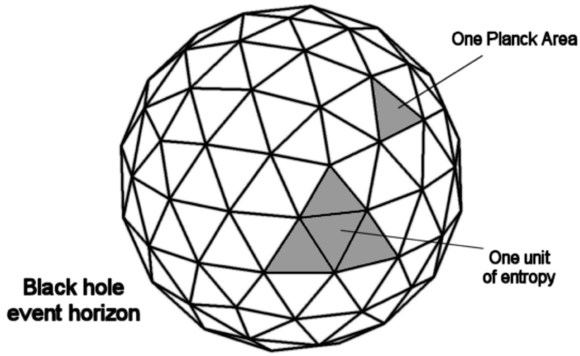


Fig. 71 BH information-entropy interpretation as a 'bit-tiled' EH surface.

So, when someone throws a chunk of matter with some amount of information-entropy into a BH, the amount of matter grows, which means that the area of its EH also grows, increasing the information-entropy storage capacity of the BH itself, represented by the number of bits which cover that surface. Therefore, the information content of something falling into a BH is not lost but, instead, gets stored onto the area of the EH. This area is no longer something we must think of as being just a surface in a 3D space; rather, it is a holographic projection of a piece of a boundary at infinity of a 2D AdS universe without gravity into a 3D space-time with gravity.

So, overall, putting the BH complementarity principle together with the holographic principle, the BH information paradox seems to find a resolution. However, many questions remain open which cast a shadow onto this approach. This increases a sense of incertitude – in many, also a strong skepticism, if not even open opposition.

First of all, as already pointed out, the AdS space is not the representation of the real universe in which we live. The AdS space-time has negative curvature – that is, it represents a universe where the gravitational forces are attractive and tend to decelerate the cosmic expansion. Whereas, the real universe has a positive curvature (this, at least, according to the current interpretation of the observational data), which means it is subjected to a still-unidentified repulsive force generated by the dark energy, which causes the cosmic expansion to accelerate. It is unclear whether the real model of space could solve the BH information paradox as the holographic principle does in the AdS context.

Then, the attentive reader might have noticed that the BH complementarity and holography help us only to a certain point. The fact is that, because BHs evaporate, they will radiate away all the BH's mass and, sooner or later (much, much later for solar-mass-sized BH!), nothing will be left of the BH due to Hawking radiation. The stored information on the EH will have to fly out again in the Hawking radiation but must still remain entangled with that inside the BH. If someone waits long enough, one will observe a BH shrinking in size and finally disappearing, leaving only a gas of quantum particles in a mixed state from which no information will be gained – not even in principle – about the previous state of matter that initially built up the BH. Therefore, up to a certain time, any information about a physical object falling into a BH is lost forever anyway. The black hole complementarity becomes inconsistent and holography won't save us from contradictions, either.

This can also be seen from the area-entropy relationship. Due to the evaporation process of BHs, the area of the EH must decrease in time. This requires that the information-entropy content must do the same. However, this contradicts the idea that, if the particles behind the horizon are supposed to be entangled with the emitted radiated ones (the overall quantum states of the BH plus its Hawking radiation is in a pure state), the entropy of the BH should, instead, remain the same as that of the outgoing radiation. It can be shown that the contradiction must show up, at the latest, at a halfway point of the evaporation process, when half of the information is emitted, and whereby there can no longer be sufficient information content on the horizon surface for holography to represent the BH quantum state. The entanglement of the BH with the Hawking radiation photons can continue only up to a certain point in time, from which onwards the shrinking surface area of the EH will have become too small to contain all the information of the interior of the BH. After that point, if the theory wants to save its consistency, it must postulate that the photons emitted should also be entangled with the radiation emitted in previous times. The photons should now be entangled with the radiation emitted previously and with the BH. That is to say, an outgoing particle is entangled with another two particles – namely, the one inside the BH, with another Hawking radiation photon radiated in the past, but without being the system of all three particles in a pure state. This is, however, not allowed by QM, which forbids a quantum bipartite entangled system from becoming entangled with another independent system, becoming, thus, a tripartite entangled one. This is what physicists call a violation of *'monogamous quantum entanglement'*. Such a type of conversion violates unitarity and the equivalence principle as well and we find ourselves at the starting point again!

If that kind of forbidden 'quantum conversion' would effectively take place near the EH of a BH, it would mean that entanglement between the infalling and outgoing particle gets broken and, as a consequence, a huge amount of energy would be released, creating a sort of 'firewall' which would almost instantly incinerate everything which would pass through the EH, contrary to what the equivalence principle states. Monogamy of entanglement stands in the way of the BH complementarity, suggesting the possible existence of a firewall near the EH. This is called the *'firewall paradox'*, which implies that the 'no-drama scenario' is quite dramatic: a bold (or stupid) astronaut falling freely into a BH would not only be stretched and torn apart into thousands of pieces due to gigantic tidal gravitational forces but would also be quickly fried up while nearing the EH.

We have reached the end of this section on QM and information theory with not many certitudes, as sound science should offer us. What all this seems to suggest is that some of our foundations in QP and/or GR must be given up when we consider physics in extreme regimes like that of a BH. The equivalence principle of GR might no longer hold, or unitarity might be violated in some circumstances, or our understanding of what quantum entanglement is must be reviewed or, perhaps, for whatever reason we actually don't know, Hawking radiation just doesn't exist, though it arises as a natural consequence of the former principles. Nobody knows and that's where the modern attempts to unify QM with GR are at this stage. What we end up with is that, for each theory that tries to marry QM with GR and that is supposed to solve a problem, another problem pops up, which leads to other paradoxes or inconsistencies.

If you feel comfortable with all this, fine. Otherwise, if you feel confused, don't bother. Many physicists would tell you "shut up and calculate!" until you find the final theory that saves us all. However, if you do not subscribe to this working philosophy, be assured that you are not alone. Many high-ranking physicists do not embrace all this without skepticism; some take a strong critical stance. We must take these theories and the current developments in theoretical physics with a grain of salt, to say the least. As long as there is no experimental evidence that supports them, ST, the holographic principle, the firewall paradox, and even Hawking radiation must be considered unverified theories. So far, there is no evidence supporting them and nothing indicates that there will be any time soon. Many have the feeling that the modern foundations of physics lost itself in a quantum quagmire. Something has happened to theoretical physics since the times when Planck, Heisenberg, Schrödinger, Pauli, Dirac, and many others developed relatively simple and neat principles or formulas that made testable predictions and, in most cases, could not only be soon effectively dismissed or verified experimentally but could explain lots of things that previously could not be explained otherwise. Nowadays, instead, we find ourselves overwhelmed by a plethora of theories, conjectures, and sometimes wild speculations that have led to paradoxes or inconsistencies. Most of them arise from very complicated and elaborate mathematical formulations and no longer have anything of the simplicity and elegance of, for example, a principle of uncertainty or a Pauli exclusion principle. At the end of the day, they do not predict anything new or they predict only things that are almost impossible to verify.

The reason why we dwelled here a bit on these highly speculative and messy intellectual mumblings is to furnish you with an up-to-date overview of the mainstream theories on QG and to make it clear where modern theoretical physics actually stands. These are the topics that theorists are nowadays discussing when trying to unify the SM of particle physics with gravity in a unified frame of a theory of QG. I hope it could be useful for gaining at least an intuitive idea of what they are talking about.

V. Other contemporary quantum themes

While the previous sections could embrace different topics, placing things into one context, there are other interesting lines of research or theoretical speculations which must be addressed separately. Contrary to the uncertain fate of QC and QG, and not to mention the interpretations of QM, there have been some groundbreaking experimental realizations such as the Bose-Einstein condensates (BEC) as well as theoretical advances in cosmology such as nucleosynthesis after the first phases of the Big Bang (though, it must be said, the latter dates back mostly to the second half of the past century). Other aspects of cosmology remain highly speculative, including inflation theory or the supposed role of QM in biological processes such as in the brain as a substrate of conscious experience. In this concluding section, we will briefly review these topics, as some of them represent a specific cornerstone of modern theoretical physics or cosmology, while others are still a matter of debate and speculation among physicists and philosophers. A concluding chapter will follow, on the sociologically interesting clash between pseudo-scientific beliefs and scientism.

1. Bose-Einstein condensates

In Vol. I we extensively discussed the difference between bosons and fermions. So far, we have considered only particles – that is, the integer-spin bosons and the half-integer-spin fermions particles. We know the photon as the typical example of a boson, namely, a spin $\pm\hbar$ particle mediator of the EM force, while protons, neutrons, and electrons are the $\pm\frac{\hbar}{2}$ spin fermions. We also pointed out that the peculiar properties of ordinary matter, such as its solidity or resistance to penetrability, are due to, among other factors, its composition of fermions which are described by an anti-symmetric wavefunction and which must obey the Pauli exclusion principle. This disallows fermions from occupying the same quantum state – that is, they can't be in the same place with the same energy level at the same time. It is this which allows the universe to come into existence in the form of solid, liquid, and gaseous substances made of atoms, the physical properties of which can be traced back to the table of elements and molecules that bind it together. Meanwhile, photons, being boson-type particles with symmetric wavefunctions, have no such physical limitations imposed by the exclusion principle. They can occupy the same quantum state – that is, overlap and not interact with and among each other (apart from very special and extreme conditions in high energy regimes). Several photons of different energies can be in the same place at the same time.

Bosons, however, must not necessarily be force mediators or have only unit integer spin. They don't even necessarily have to be particles. Bosons are, per definition, all integer-spin objects and, therefore, can also be atoms or molecules or even larger structures, provided that their net spin is an integer of the reduced Planck's constant. If an atom is electrically neutral, the number of protons and electrons is the same and their half-integer spin cancels. Therefore, not protons and electrons determine the net spin of an electrically neutral atom but neutrons do – that is, if the atom is a *'composite boson or fermion'.* If the number of neutrons piling up in the nucleus energy configuration (always having to follow Pauli's exclusion principle) is even (odd), they sum up into an integer-spin object (half-integer object) and the atom is a composite boson (composite fermion).

Let us examine some examples of bosonic atoms. Hydrogen is the first and most simple composite boson (the non-isotopic ^{1}H with zero neutrons in its nucleus) but its electron magnetic moment pairs with that of the proton's and is, therefore, also magnetically neutral – a property which makes it difficult to control and to contain inside a magnetic field for experimental purposes. (Recall how the magnetic moment of electrons was used in the SG-experiment.) The simplest non-reactive example is the Helium-4 (^{4}He) atom, with two protons and two neutrons in the nucleus and the two electrons in its atomic shell. Because of this configuration ^{4}He is an inert element and, being the lightest noble gas, has extremely low melting and boiling points (about 1 and 4 Kelvin, respectively, at 25 bar pressure). However, again, there is no net magnetic moment (neutrons, being electrically neutral, can't build up a magnetic momentum) but, as we shall see, this is by far not its only peculiar property. Good candidates for experimental usefulness are, instead, the alkali metal isotopes with an even number of neutrons – that is, are bosonic elements of the periodic table which, however, have an unpaired electron in the outermost shell and which give them a magnetic moment.

The question is: can also composite bosonic atoms or molecules all be in the same energy state at the same time, provided the temperature allows for this to happen? This is unlike fermions, which must follow Pauli's exclusion principle. In fact, as long as a gas of bosons is subjected to a warm environment, all the atoms have different energy states. The thermal excitation scatters them around and, due to the inelastic scattering processes, they absorb and continuously emit energy quanta and each atom has its own energy. However, if one cools them down to cryogenic temperatures at almost the *'absolute zero temperature'* (0 K, -273.15 °C, -459.65 °F), the bosonic atoms all tend to occupy the same energy level – namely, the ground state energy level, just as the electron in the hydrogen atom has a ground state which is ultimately determined only by Heisenberg's uncertainty. This implies that when no thermal effects from the environment remain, while fermions nevertheless must pile up in spin-up and spin-down pairs on each energy level (called the *'Fermi energy'* levels E_f) forming a *'Fermi sea'*, bosons instead are free to exist all together in the ground state at the same time, as illustrated in Fig. 72. (Compare this to the last figure of the bosons, fermions, and Pauli's exclusion principle chapter in Vol. I.)

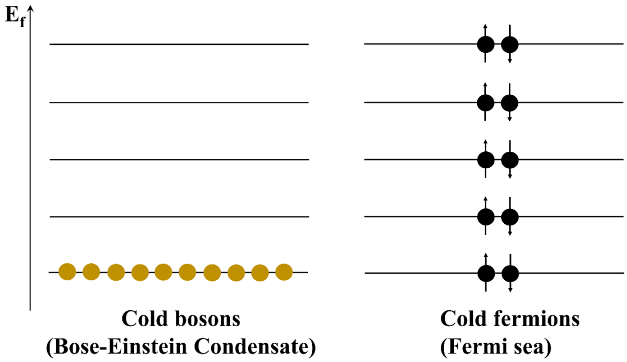


Fig. 72 The different 'condensation forms' for a Boson- and Fermi-gas.

Whereas, we know that for particles, atoms, molecules, or any physical quantum object, the de Broglie relation associates it with a wavelength ($\lambda = \frac{h}{p}$). The lower the temperature, the smaller the momentum (the velocity) of the atoms and the larger their wavelength. Meanwhile, even at only a few degrees Kelvin above absolute zero, the atoms have relatively high impulses and are localized by the de Broglie wavelength. (See Fig. 73 left.)

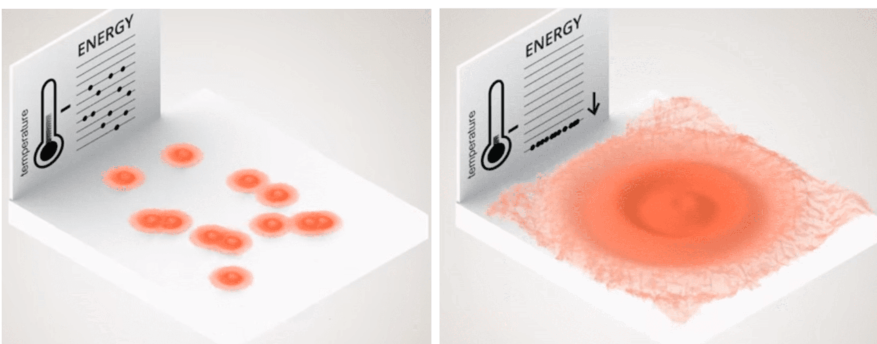


Fig. 73 The coalescence process of bosons into a Bose-Einstein condensate.

However, when the gas is cooled down to a sufficiently low temperature, the wave-packets spread out and overlap each other until they form a unique, undivided gas of atoms, all in the same quantum state and described by a single wave-packet. (See Fig. 73 right.) Therefore, they will not only group themselves into the same energy level but also coalesce into a unique and almost-macroscopic quantum object that can be described by a single wavefunction or matter-wave. All this implies that, in this extreme physical state, the atomic bosons do not become a solid but, rather, condense and form a new state of matter (beside that of the solid, liquid, and gas state), called a *'Bose-Einstein condensate'* (BEC).

This new state of matter was predicted in 1924 by Einstein, who was inspired by a novel statistical description introduced by the Indian physicist Satyendra Nath Bose, who, as we already know, is the guy to whom bosons owe their name. Bose first recognized how bosonic particles must obey a different energy level quantum statistical distribution than that of fermions. He applied this to photons and re-derived Planck's blackbody radiation in a different manner than Planck did. His approach was yet another theoretical validation of Planck's work on the blackbody spectrum.

Fig. 74 Satyendra Nath Bose (1894-1974).

Initially, Bose wrote a paper explaining his theory and sent it to a physics journal, which rejected it. He then asked Einstein for help, stating humbly but self-assuredly: *"Though a complete stranger to you, I do not feel any hesitation in making such a request. Because we are all your pupils though profiting only by your teachings through your writings."* [24] Einstein was impressed; he translated his paper into German and published it in a German journal under Bose's name (an internal social-political academic dynamic that speaks volumes about the lack of acceptance and recognition of outsiders in the academic community, both then and nowadays).

Shortly thereafter, Einstein extended Bose's theory to atoms. What distinguishes the statistics of integer spin particles from those with half-integer spins is not only the fact that the former can attain the same quantum state but also that it describes the distribution of identical particles. Fermions in the same quantum state are still distinguishable by their opposite spin; bosons, however, are indistinguishable in terms of the strong quantum-mechanical sense we outlined for the principle of quantum indistinguishability.

This changes everything regarding the number of possible quantum microstates that a set of bosons can acquire. To get an intuitive understanding of what this is about, think, for example, of two distinguishable particles with two possible quantum states, $|0\rangle$ and $|1\rangle$. The system can attain four possible configurations – namely, $|0\rangle|0\rangle$, $|0\rangle|1\rangle$, $|1\rangle|0\rangle$, and $|1\rangle|1\rangle$. However, if the particles are indistinguishable in the strict quantum mechanical sense, then the second and third quantum states must be considered one and the same. Therefore, the system can acquire only three different configurations. Extended to a huge number of particles, this leads to a completely different number of possible microstates – that is, a different energy level statistical distribution for bosons compared to fermions. From there comes the name of *'Bose-Einstein statistics'*, valid for bosons on one side, and the *'Fermi- statistics'* of fermions on the other side.

The problem with BEC at the time was, however, not theoretical but only practical. It took 71 years to prove that Bose and Einstein were right because it is quite difficult to create perfect vacuum chambers cooled down at a few billions of degrees Kelvin and perfectly shielded from any outside disturbance to avoid the loss of quantum coherence. In 1995, however, the technology was mature enough to create such extreme physical conditions. Eric Cornell and Carl Wieman of the University of Colorado at Boulder were able, by a combination of laser cooling and magnetic confinement mechanisms, to create the first BEC in a laboratory.

Laser cooling is a technique in which atom or molecule samples can be cooled down to near absolute zero. If the direction and wavelength of the laser photons are tailored carefully, they can be absorbed and reemitted many times by the atoms in such a manner that they are slowed down – that is, their thermal kinetic energy is reduced to almost zero. Magnetic trap confinement is, of course, necessary not only to keep the gas in place but also to prevent it from literally falling to the surface due to gravitation and avoiding the interaction with the container's wall. (This is why physicists are eager to reproduce the same experiment in the absence of gravity in space, as was done recently on the International Space Station.) With laser techniques, also, the temperature of the BEC can be measured (due to the Doppler effect – that is, the change of the wavelength of the reflected light, which conveys information about the speed of the atoms) or even photographed! The latter method, being invasive, obviously destroys the extremely fragile BEC but is, nevertheless, fast enough to convey the necessary readout. For their achievements, Cornell and Wieman, together with Wolfgang Ketterle at MIT, who confirmed their discovery and further enhanced their technique, received the 2001 Nobel Prize in physics.

Because a BEC has all the atoms in the ground state, the transition between a 'warm' bose-gas to a BEC is, therefore, reflected in the velocity distribution of the bosons making up the gas. Before the condensation, they behave like all other normal gases – that is, the atoms scatter around and are subjected to Brownian motion. However, because the gas has become a BEC, the single constituents coalesce and become a unique and undivided matter-wave described by a single wavefunction for all the particles. To be more precise, the single quantum 'super-particle' that the gas has become will display only one thermal velocity distribution, with the important caveat that it is still somewhat uncertain due to Heisenberg's uncertainty.

Fig. 75 shows the spatial atom-cloud density profile of a gas of 2000 rubidium atoms (the isotope ^{87}Rb) before, during, and after the BEC phase transition. The height at a certain point gives the density while the colour gives the velocity. Red is the 'hottest' while white is the coldest, so to speak (for the B/W edition: from the outside to the center of the peaks, the colours are as follows: red at the bottom, then yellow, green, blue, and white at the peak, indicating a decreasing velocity – that is, decreasing kinetic thermal energy).

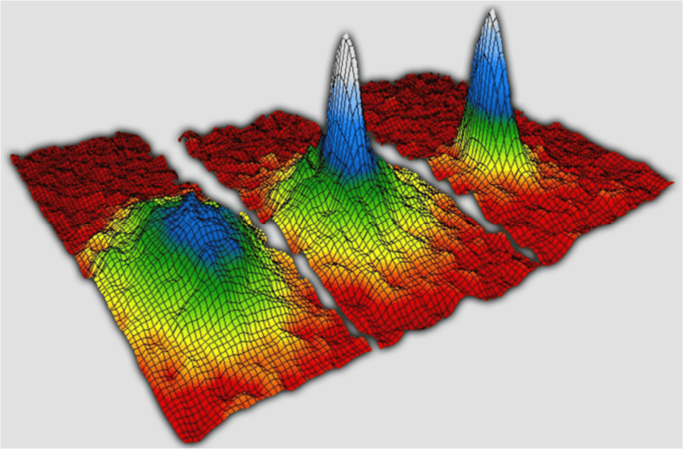


Fig. 75 Bose-Einstein condensate velocity distribution. [25]

In the left plot, at about 200 nK, no BEC took place. The atom's distribution is broad and no 'cold' atom in its ground state is observable. The middle plot (100 nK) displays the gas just after the phase transition to a BEC. The white-coloured top of the peak reveals the existence of atoms in the ground state. The third picture (almost 0 nK) shows the fully realized coalesced pure BEC with a highly localized peak where a large fraction of atoms occupies the ground state. The latter thermal distribution contains many atoms in the ground state but still with a finite breadth which is determined by the uncertainty principle.

Therefore, as weird as this might seem, composite bosonic atoms or molecules can also be in the same place at the same time, even if they are made of the proton, neutrons, and electrons which are fermion particles.

Another interesting aspect is that, because a BEC is a matter-wave described by a single wavefunction, it is natural to expect that two BECs traveling through each other behave just like two light wavefronts. Despite being a gas made of several thousands of atoms that we imagine as separate entities, when a BEC encounters another BEC, the typical interference phenomena must manifest, such as those we have discussed extensively for EM waves. This implies that when we overlap two clouds of BEC atoms, their matter densities do not simply sum up but, rather, interfere with each other, creating a spatial density profile with periodically arranged interference maxima and minima analogous to two interfering light beams which produce a similar scheme on a screen. Fig. 76 shows an image taken in 1995 by Ketterle's group; the interference fringes between two freely expanding sodium atom BECs are separated by two 40 ms time frames. (Compare this with the figure of the superposition of two waves and the interference pattern at an inserted screen in the chapter on bosons, fermions, and Pauli's exclusion principle of Vol. I.)

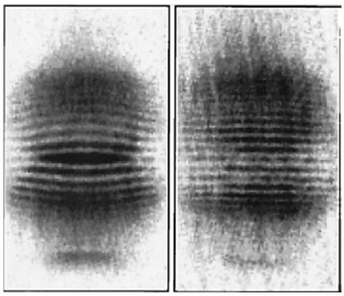


Fig. 76 Two BECs interfering with each other. [26]

The fringes signal a high probability of localizing a single atom; the absence of it indicates zero or almost no probability of detecting the atom in the specified location. This is a very peculiar property of quantum matter and one that we never observe in our daily macroscopic life. However, this was not news for a single or a few particles; rather, it was the first demonstration of how a quasi-macroscopic object (40 micrometers) made of thousands of atoms can display the same quantum properties as well.

Another important effect that can be observed in BECs is quantum mechanical tunneling. We know what that means for a single particle: It can 'tunnel' through a potential barrier even if, classically, it does not have enough energy to do so. Nothing in QM prevents us from extending this weird quantum property of particles to atoms, molecules, or an entire gas. In fact, a macroscopic quantum many-body tunneling could be observed; at least part of a BEC can, indeed, tunnel a barrier that a classical particle could not overcome.

Later, it was found that, in an analogy to BECs, fermions can also produce so-called *'fermion condensates'*. In fact, provided that particular physical conditions are met, fermions can pair to form an integer-spin molecular boson, such as with two ^{3}He atoms. Fermionic condensates have even lower transition temperatures.

Without going into further detail (which would lead us into complicated considerations of statistical physics), as a concluding side note, it is worth mentioning some other weird properties of BECs. For example, ^{4}He is a bosonic atom with the very interesting property that, once cooled down to a BEC, it displays *'superfluidity'*. A superfluid is characterized by zero viscosity, i.e., the ability to flow without friction, a property which leads to quite surprising effects, such as a fluid climbing the walls of a container or displaying vortices that never stop spinning. BEC and superfluidity are closely related but are not necessarily the same. A BEC system does not necessarily exhibit superfluidity and vice versa. Moreover, BECs can be superconducting – that is, become electric conductors with zero electric resistivity. All these properties arise because BECs are a new state of matter. An almost incredible property of a bosonic gas that arises as a direct consequence of single atoms having lost their individuality is that its optical refractive index tends to infinity. The refractive index of a material determines how fast light travels through an optical medium. The larger the index, the slower the light will travel through a transparent medium. If an object has an almost infinite refractive index, it has the virtual ability to 'freeze' light. In fact, this is a very strange property that has been proven to be true.

In Nature, we observe the solid, liquid, and gaseous state. We also know a fourth state to exist – namely, *'plasma'*, which is a 'gas' of particles, electrons, and protons no longer bound to a nucleus. Plasma is known to exist in very hot places like the Sun. The BEC can be considered a fifth aggregate state of matter. Moreover, if BECs exist only in extremely cold or extreme environments, an open question still remains. According to some theoretical calculations, it might exist in the interior of neutron stars, the remnants of old dead stars, which, compressing matter to enormous pressures, might eventually be able to bind fermions into bosonic composites.

However, apart from these theoretically interesting quantum properties, BECs might one day also find some practical applications. Because they are quantum macroscopic objects, they might turn out to be useful in building quantum circuits for QC. A little cloud of BEC atoms could, in principle, be used as a single qubit which will be less vulnerable to noise and interference from the environment, to which single-particle qubits are so sensitive. Additionally, its optical properties could potentially lead to new applications in light-based telecommunications.

The research on BECs was widely in fashion and made the greatest advances during the first decade of the millennium. Nowadays, their popularity has faded a bit in the shadow of QC. However, whenever its application for practical purposes might become a reality, or even if not, the fact that it is possible to prove the existence of quantum macroscopic phenomena is a milestone in physical science. This allows us to probe the bizarre quantum realm on a large scale instead of having to probe single particles. One wonders what Bose and Einstein would have said if they could have seen their brainchild become reality.

2. Big Bang nucleosynthesis

The reader may already know that modern cosmology tells us that the universe began with a huge explosion, the famous *'Big Bang'*, 13.7 billion years ago. According to this theory, in the past, all matter and energy were concentrated in a much smaller volume of space. Nowadays, we have strong evidence that the Universe is expanding; in fact, this expansion is accelerating. In the first phase of the universe's existence, matter and energy were subjected to such extreme pressures and temperatures that we can state with certainty that the laws of classical physics didn't hold. To properly understand the physical conditions of the primordial universe and the phenomenon that led to its origin, the laws of quantum physics are necessary for investigating the creation of energy and matter in the first place and nuclear physics to know how the first elements came into being. The former approach is that of quantum cosmology, while the latter developed into the theory of Big Bang nucleosynthesis. This is what we will discuss first in this chapter.

Because the SM of particle physics is a rock-solid theory which could not be superseded by a theory of QG, likewise, while present quantum cosmology remains speculative, the foundations of modern cosmology are still rooted in very successful cosmological basics which were developed in the 1950s, on top of the discoveries of quantum physics and nuclear physics. Particularly successful in the microphysical domain became the theory of *'Big Bang nucleosynthesis'*, a well-established theoretical framework which describes the creation of the first light elements and which became the primordial bricks from which all the elements of the periodic table come. It received strong observational validation.

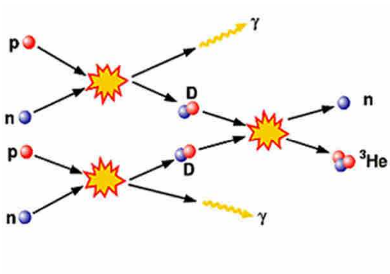
Nowadays, we know that most of the elements that built up the stuff of which we are made were not synthesized during the Big Bang phase. Rather, they were synthesized in the stellar interior through nuclear fusion processes. However, stars could come into existence only because of the elements created during the early phases of the Universe – first and foremost, hydrogen and helium and, to a much lesser degree, lithium and beryllium.

In the next chapter, we will discuss what happened after the first 10^{-32} seconds after the cosmological inflation phase – that is, what happened almost instantly after the universe's creation can nowadays be described quite well with current nuclear physics, without resorting to QG. This is also facilitated by investigation with modern particle accelerators that can recreate, with few particles, the temperatures and pressures to which matter was subjected until about 10^{-12}s after the Big Bang. That's another reason why particle accelerators are considered such important experimental tools. When particles are smashed against each other, they allow us to literally recreate 'small bangs'. Moreover, a lot of evidence coming from astronomical observations made with powerful telescopes (on the surface as in space) has allowed us to collect a set of data which confirmed the theoretical predictions.

Immediately after the Big Bang, the universe was still too hot to allow for the formation of not only atoms but also particles such as protons and neutrons. The universe must have still been a 'soup' of even more fundamental constituents of matter – namely, a *'quark-gluon plasma'*. However, things changed quickly at those times: After about one microsecond, the universe cooled down sufficiently to allow the quarks to bind together and form protons and neutrons (and other baryons – that is, as you might recall, particles made up of two or three quarks, which, however, decay quickly in protons or neutrons themselves). However, the temperature was still much too high to go beyond that; still no atoms could form. It took only another hundredths of a second of the universe's expansion to diminish the pressure and temperatures to about 10^9 K, which led to the formation of the first elements. At that point in time, the protons and neutrons (electrons, photons, and neutrinos as well, which we will not consider here) could form the first nuclei of the light elements through nuclear fusion reactions.

Fig. 77 provides a simple scheme of how the primordial elements were formed. This simplified Big Bang nucleosynthesis scheme can give you a glimpse at how the stuff that makes up stars was built from the bottom-up.

The positively charged proton is labeled as p^+, the neutron as n^0, and the high-energy EM gamma ray as γ. The standard symbols for the nuclides (the element's nucleus) are as follows. In general, these are labeled as ${}^A_ZE$, with E being the capital-letter symbol that stands for the nuclide, A being the atomic mass number (the total number of protons and neutrons), and Z being the atomic number (the number of protons). Therefore, the difference A-Z gives the number of neutrons. The elements listed here are: H = hydrogen, He = helium, D = deuteron or deuterium, T = tritium, Li = lithium, and Be = beryllium. For example, ${}^2_1D$ is an atomic nucleus with one proton and one neutron – that is, a hydrogen nucleus with one neutron more, which is called an *'isotope'* of hydrogen. (That's why you might also find it in the literature as ${}^2_1H$.) Another isotope of hydrogen is tritium ${}^3_1T$ (or ${}^3_1H$), which is, again, the hydrogen nucleus but with two neutrons. ${}^7_4Be$ is beryllium with four protons plus three neutrons, and so on. Keep in mind that not all of these elements are stable. (The listed tritium and beryllium isotopes aren't.)



$$^{7}_{4}\mathrm{Be} + \mathrm{n}^0 \longrightarrow ^{7}_{3}\mathrm{Li} + \mathrm{p}^+$$
$$^{3}_{2}\mathrm{He} + ^{4}_{2}\mathrm{He} \longrightarrow ^{7}_{4}\mathrm{Be} + \gamma$$
$$^{7}_{3}\mathrm{Li} + \mathrm{p}^+ \longrightarrow ^{4}_{2}\mathrm{He} + ^{4}_{2}\mathrm{He}$$
$$^{3}_{2}\mathrm{He} + ^{4}_{2}\mathrm{He} \longrightarrow ^{7}_{3}\mathrm{Li} + \gamma$$
$$^{3}_{2}\mathrm{He} + ^{2}_{1}\mathrm{D} \longrightarrow ^{4}_{2}\mathrm{He} + \mathrm{p}^+$$
$$^{3}_{1}\mathrm{T} + ^{2}_{1}\mathrm{D} \longrightarrow ^{4}_{2}\mathrm{He} + \mathrm{n}^0$$
$$^{3}_{2}\mathrm{He} + \mathrm{n}^0 \longrightarrow ^{3}_{1}\mathrm{T} + \mathrm{p}^+$$
$$^{2}_{1}\mathrm{D} + ^{2}_{1}\mathrm{D} \longrightarrow ^{3}_{1}\mathrm{T} + \mathrm{p}^+$$
$$^{2}_{1}\mathrm{D} + ^{2}_{1}\mathrm{D} \longrightarrow ^{3}_{2}\mathrm{He} + \mathrm{n}^0$$
$$^{2}_{1}\mathrm{D} + \mathrm{p}^+ \longrightarrow ^{3}_{2}\mathrm{He} + \gamma$$
$$\mathrm{p}^+ + \mathrm{n}^0 \longrightarrow ^{2}_{1}\mathrm{D} + \gamma$$
$$\mathrm{n}^0 \longrightarrow \mathrm{p}^+ + \mathrm{e}^- + \bar{\nu}_e$$

Fig. 77 Left: Pictorial representation of the combined second and fourth reactions. Right: A bottom-up Big Bang nucleosynthesis scheme.

After some time, they might decay into other nuclei. For example, we can no longer observe $^{7}_{4}Be$ because it has a mean lifetime of only about 53 days (it decays into $^{3}_{2}He$) and after 13.7 billion years, nothing is left.

Following Fig. 77 from the bottom-up, one can see how the first reaction describes the decay of a neutron n^0 into a proton p^+ and an electron e^- with the emission of a type of neutrino $\bar{\nu}_e$ (an 'electron neutrino'). The second reaction tells us about the nuclear fusion of one proton p^+ with one neutron n^0, resulting in the nucleation of one deuteron nucleus $^{2}_{1}D$ plus the emission of energy in the form of a gamma ray γ. Then, as shown in the third reaction line, these deuterium nuclei $^{2}_{1}D$ fly around in the primordial universe and quickly encountered another proton p^+, which leads to a fusion process that produces $^{3}_{2}He$ plus another gamma ray γ, thereby releasing energy. The same $^{3}_{2}He$ could also be obtained via another reaction, that of the fourth reaction line, with the fusion of two deuterons $^{2}_{1}D$ plus the release of energetic neutrons n^0. However, there is a certain quantum chance that the reaction of the fifth line takes place – that is, the fusion of two deuterons $^{2}_{1}D$ could lead to the creation of tritium $^{3}_{1}T$ instead. The author leaves it to you to interpret the rest of the reaction ladder and see how it leads to the production of the other light elements.

This nucleosynthesis lasted for only three minutes after the Big Bang, which was enough time to produce mostly deuterium, $^{4}_{2}He$, with small amounts of ^3_2He and ^7_2Li (and beryllium, which isn't counted here because it was destined to 'disappear' anyway due to radioactive decay).

This is what, in its essence, the Big Bang nucleosynthesis predicts. It is the result of a study first pioneered in 1948 by the Russian-American cosmologist George Gamow. He was also the first to predict the existence of a CMB with a blackbody spectrum. (His guess was 5 K temperature, against the observed 2.7 K.) However, Gamow did not realize that all the elements could be produced in the early universe. His co-workers, Ralph Alpher and Robert Herman, completed the theory by pointing out how there couldn't be enough time to synthesize heavier elements. These could come only afterward by means of nuclear fusion reactions inside the stars. That's why you might have heard scientists say that we are 'stardust'. The elements heavier than lithium of which our bodies are made (especially carbon, oxygen, iron, and other building blocks essential for the emergence of life) were synthesized in a stellar furnace billions of years ago by a progenitor of our Sun and then were ejected into the solar system through a supernovae explosion.

Fig. 78 George Gamow (1904-1968).

The Big Bang nucleosynthesis is not just a theory. Nowadays, it is backed by solid observational evidence. In fact, it makes a clear cut predictions which can be tested experimentally – namely, that once the density of matter is known, the model fixes a precise abundance of deuterium, helium, and lithium relative to hydrogen. We already knew of these relative abundances from astronomical observations coupled with spectroscopic analysis of the light coming from throughout the universe, which indirectly furnished the universal density of matter and radiation. The matter density of the universe, being defined as the ratio between the number of all baryons (essentially, protons plus neutrons) and the number of photons, could be measured directly. In fact, this is what the Wilkinson Microwave Anisotropy Probe (WMAP) satellite did by measuring the CMB. The CMB that we can observe in the sky today is the microwave radiation left from the Big Bang 'echo' corresponding to the *'recombination epoch'*, in which the electrons began to combine with the nuclei synthesized earlier and created the lighter chemical element atoms. This recombination epoch dates back to about 380,000 years after the Big Bang. The analysis of the CMB, such as the density of radiation, also furnishes the density of matter. WMAP furnished a value of a ratio of about 5×10^{-10} baryon to photon (the vertical line in Fig. 79 – the axes are drawn in logarithmic scale, see Appendix A Ib).

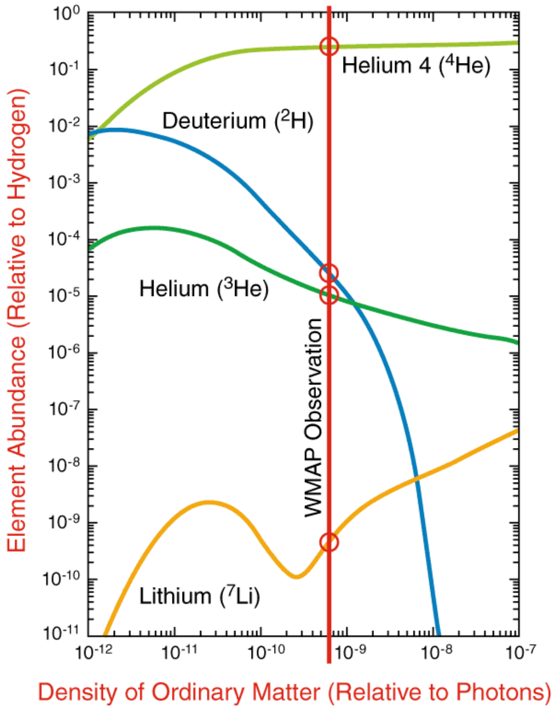


Fig. 79 Element abundance graph. [27]

Therefore, on average, there are two billion photons for every proton or neutron in the universe. Plugging this into the Big Bang nucleosynthesis theory (the curves in Fig. 79), the relative light element abundances are fixed. It turns out that under these conditions, the abundance of ^4_2He is roughly 24% of hydrogen. As little as 0.01% is left for deuterium, one part over 10^5 for ^3_2He, and about one part over 5×10^9 for lithium. These relative abundances are the circles in the graph. One can see how precise the match is between the theory, the observed abundances, and the observed matter density: For three of the four elements, the circles match almost exactly the theoretical predicted relative abundances with the observed abundances. Only lithium is rare and, on a logarithmic scale, it is not apparent how the theory predicts a three- to four-times-greater amount of lithium than actual observations show. It is called the *'cosmological lithium problem'* but it is one of the last inconsistencies of the current theory. For all other elements produced shortly after the Big Bang, the theory conforms well to the measured data.

This neat correspondence between observational data and theory (together with the neat match we could observe with the black body spectrum of the CMB at the end of the chapter on the blackbody radiation in

Vol. I) is yet another triumph of the Big Bang model and assures scientists that the basic mechanisms with which elements were synthesized during the first three minutes after the Big Bang theory are understood. Moreover, it is another nice example of the history of science in which particle physicists studying the microcosm collaborated successfully with astrophysicists investigating the macrocosm. The history of the universe could be graphically summed up as shown in Fig. 80.

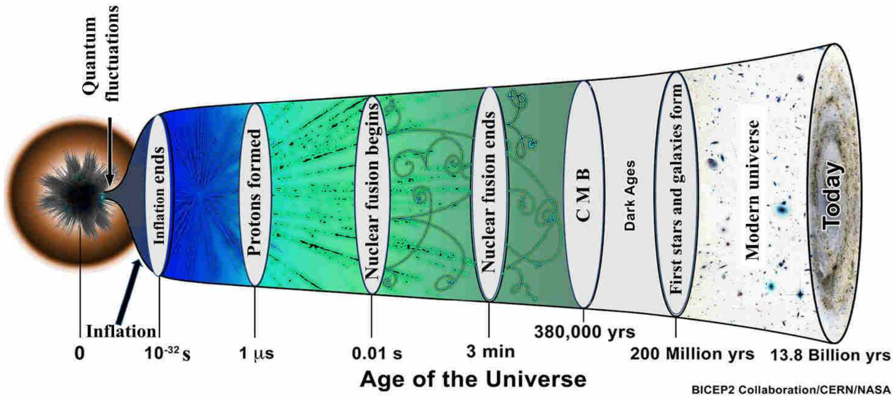


Fig. 80 The creation and expansion of the universe throughout its different epochs.

During the time between 380,000 years and 200 million years after the Big Bang, the universe must have been a boring and dark place, as photons were all high-energy light particles invisible to the human eye. However, the temperature after recombination was still about 3000 K. Only at the end of these 200 million years did the universe cool down sufficiently to form giant hydrogen clouds, which collapsed due to the gravitational pull forming galaxy clusters, galaxies, and stars and where, later on, planets could form.

3. Quantum cosmology and cosmic inflation

Let's now turn our attention to modern quantum cosmology. As you might expect, besides QG, this is still an uncertain and speculative field of research which cannot be summarized in a few pages. We can, however, take a quick glimpse at one of its main representative theories, which is *'inflation cosmology'*.

Apart from the fact that astronomical observations show how galaxies fly apart, indicating that, in the past, the universe must have been much smaller and with a common origin, there is a very simple reason why the Universe must have been much smaller than it is in its present state: its isotropic and homogeneous distribution of matter and radiation. The almost perfectly uniform CMB radiation and its related Planckian black body spectrum (recall the end of the chapter on blackbody radiation in Vol. I) show that it changes only by one part over 100,000 all over the sky. Isotropic means that the distribution of galaxies on the sky always looks the same for all directions, while homogeneous means that at every cosmic distance, the density of galaxies is uniform and, on average, the same, as it is in our galactic neighborhood. In some sense, this aligns with the so-called *'cosmological principle'* already stated by Newton, which rests on the assumption that there is no reason to believe that any part of the universe should be special or have some peculiar property which should significantly differ, apart from small, random statistical fluctuations, and that, therefore, on average, the large-scale distribution of matter and energy in the universe must be expected to be homogenous and isotropic.

However, things are not as easy as that. It looks like all the different sky patches on the cosmic horizon – that is, the part of the universe we can observe as far as 13.7 billion light-years away – 'know' each other and 'agree' to take all the same appearance, despite being much too far away from each other to be able to communicate. For example, let us take two opposite patches of the sky (with us, as human observers, being in the center in between). These are 27.4 billion light-years apart and have no means by which they can be in causal contact with each other. They do not have enough time to inform each other about their physical state, considering that the fastest light signal would take 27.4 billion years to do so. This suggests that all parts of the cosmos must have had a common origin and that there must have been a time when these different regions, which we now observe on the cosmic horizon, were causally connected. A common origin where they could interact sufficiently to render the distribution of matter and energy they contained isotropic and uniform and, while time was passing, was diluted in a cosmic expansion. (As an analogy, think of an expanding gas of particles.)

However, this can't be the whole story, either. According to current observations, the rate of expansion of the Universe (one can measure the speed of the Galaxy shifting apart by observing the change in frequency, the so-called *'Doppler shift'* of the EM spectrum) was much too fast and could not allow the hot matter of the primordial universe to mix for a sufficiently long time that would have led to the almost perfect isotropic homogeneity we observe with telescopes and in the tiny ripples of the CMB today. This smooth distribution of matter and radiation in the universe, despite their regions being causally disconnected, is the *'flatness problem'* that modern cosmology had to tackle. Another physical mechanism must have conspired to flatten the universe's matter and radiation as we know it. It is here where quantum physics and the cosmic inflation theory come into the play, which was first proposed in 1979 by the American theoretical physicist Alan Guth at Cornell University.

Fig. 81 Alan Guth.

The basic concept upon which cosmic inflation rests is the hypothetical concept of a *'false vacuum'*, which describes the zero-point energy vacuum in terms of scalar fields. (Recall the chapter on the zero-point energy in Vol. I, together with the description of scalar and vector fields in chapter III.1 of the present volume.)

In the very beginning, the universe came into being from an empty region of space-time which was in a different energy configuration than that of the present vacuum. We saw that the zero-point energy that describes the 'true' vacuum is the minimum energy state of empty space. Due to the laws of QP, even a vacuum has a minimum non-zero energy ground state. (As an analogy, think of the ground state of the hydrogen atom.) An extension of this concept is the false vacuum scalar field φ (see Fig. 82 and Fig. 34 of section III.1), which can have two or more energetic minima. (Again, as an analogy, think of the hydrogen atom's different energetic configurations.) The *'true vacuum'* is in a global minimum energy state (the smallest minimum possible), while the false vacuum is in an excited energy state of a local minimum (i.e., a minimum energetic configuration but not the smallest possible one). This false vacuum region of space is 'trapped', so to speak, in a local minimum energy state and is therefore unstable, as it could 'roll down' towards the stable true vacuum's global minimum energy. This transition can occur, for example, due to quantum fluctuations which may lead to quantum tunneling from one to the other state. The probability at which the false vacuum transitions towards the true vacuum depends on the energetic height of the barrier separating the two minima. Cosmic inflation conjectures that the primordial universe was in a state of false vacuum and that the Big Bang was the effect of a *'vacuum decay'* – that is, a transition from this unstable state to the stable true vacuum state in which we now live.

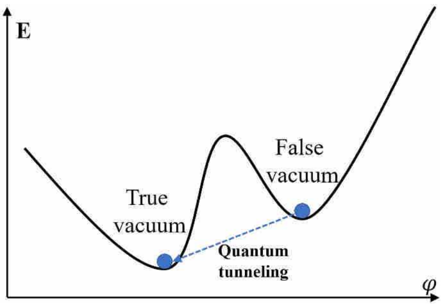


Fig. 82 Distinction between a 'true' and 'false' vacuum scalar field.

One can understand how this false vacuum region, also called the *'inflaton field'*, can initiate a phase of exponential expansion if GR is applied to QFT. According to Einstein's field equations, matter and any sort of energy density are not the only phenomena that create an attractive gravitational field; pressure does, too. Meanwhile, Guth recognized how the false vacuum has the peculiar property to give rise to negative pressure (suction) equal to its energy density and a positive repulsive gravitational energy. This repulsive gravitational field in the very first phases of the universe's existence is the driving force that led to the Big Bang. This entails that, if a patch of an early universe is in a false vacuum state, the repulsive gravitation drives it to an extremely fast expansion. The empty space of the false vacuum is 'inflated' at an exponential rate and the inflating region will enlarge by many orders of magnitude, leaving behind the non-inflating region as a remnant microscopic in comparison. Inflation theory predicts expansion factors of about 10^{25} of an initially incredibly small universe of 10^{-24}cm (the size of a proton is about a 10^{-15}m) for an almost instantaneous interval of time of about 10^{-32}seconds. The laws of physics governing the universe in these extreme conditions – for example, with temperatures up to 10^{32}K – are those of QG at the Planck scale. Therefore, we can't say much about what really took place. Rather, we can only apply GR to QFT as far as possible, as has already been done for BHs and Hawking radiation (an approach also called *'semi-classical quantum gravity'*). During this extremely tiny inflationary time interval, this 'bubble' of a false vacuum maintains a constant positive energy density. Therefore, the total energy of the expanding volume (imagine a tiny expanding sphere in which volume scales proportionally to the third power of its radius) must grow at a factor of $(10^{25})^3 = 10^{75}$.

However, this is not a violation of energy conservation. In fact, this is not really surprising, as any attractive gravitational field (the one we experience in our daily life) is already described by negative potential energy in CP, and GR makes no exception in this sense. The expanding bubble 'borrows' its positive potential energy from the gravitational field: The internal increasing gravitational repulsive positive potential energy of the nucleated bubble is compensated for by the negative potential energy of the growing attractive gravitational field. When the inflationary field relaxes to the vacuum, this inflationary epoch stops and the normal expansion settles in. The huge potential energy of the inflaton field decays into the elementary particles that filled the universe with matter and radiation. It is believed that most of these particles must have been photons; this was the radiation-dominated *'reheating phase'* of the universe.

The point is, this short exponential inflationary expanding phase was extremely fast, so that it literally 'flattened' any remaining inhomogeneity in the baby-universe. As an analogy, think of a corrugated elastic surface stretched out to an area which is several orders of magnitude larger than its original size. It is intuitive to see how any irregularity will quickly straighten to an almost perfectly flat surface. This aspect of cosmic inflation could be the reason why the causally disconnected observable universe looks so isotropic and homogeneous. The inflation phase smoothed out all the inhomogeneities and anisotropies by means of a *'metric expansion'* of a patch of space that was originally causally connected but that has now taken on cosmic dimensions. The tiny fluctuations in CMB radiation that we can observe nowadays correspond to the initial quantum fluctuations present during the inflation epoch. One might say that what we see today in the ripples of the CMB is an inflated and magnified still image of the primordial quantum fluctuations of the universe in its first phases of creation. (What a wonderful gift to astrophysicists!)

It is worth mentioning that these primordial inhomogeneities also later determined the galaxy formation and evolution which led to the present large-scale structure of the universe. These tiny perturbations departing from the average primordial fluctuations allowed for the gravitational collapse of matter and for the formation of cosmic structures such as galaxies and clusters of galaxies. Therefore, the cosmic inflation theory initially became a very popular and widely accepted theory of quantum cosmology because it provides a good answer to the flatness problem posited by the cosmological principle and yet tells us how, at a local scale, complex structures can also come into being.

However, this doesn't mean that, after the inflation epoch, the universe was a tranquil place to be in. The initial temperature of the universe during its first inflationary phase fell from 10^{27}K to 10^{22}K. This was such an extremely high temperature that the EM force and the strong and weak nuclear forces were still not three distinct interactions. Only later, when the universe cooled down further, did they split into the forces we know nowadays, due to a mechanism of QFT called *'spontaneous symmetry breaking'*. We might think of it as a similar mechanism that manifests itself when hot water vapor is cooled down to a unique pressure and temperature at which all three phases of water (vapor, ice, and liquid) are allowed to coexist (what chemists call the *'triple point'*).

All this sounds very interesting and encouraging to many astrophysicists. Cosmological inflation seems to be a plausible answer to the several open questions of modern cosmology. However, inflation theory remains a speculative theory because there is no consensus as to whether the CMB radiation we observe is, indeed, the cosmic archaeological remnant – that is, the map of the primordial quantum fluctuations. Moreover, like string theory, cosmic inflation is not a unique theory but, rather, contains many different models that make different predictions. To obtain the universe we observe, the initial inflaton scalar field must have had some properties, such as a specific magnitude and frequency of fluctuations. The only scalar field known so far is that which describes the Higgs boson discovered in 2012 but experimental data at the LHC seems to disconfirm this identification as the field at the origin of the universe. Any other initial conditions of the false vacuum would have produced an entirely different universe which, eventually, couldn't even have given rise to stars and galaxies – and, therefore, not even to life. In fact, a slightly different inflaton would have led to a universe that would have collapsed much too early after its creation to form stars and habitable planets. Or, vice versa, it could have produced much-too-high density concentrations of matter that would have collapsed into giant BHs, rendering the rest of the universe a dark and cold place without light and life. What caused exactly these initial conditions to 'fine-tune' the universe in which we live, and allowed us to contemplate and study it, is not known. This is yet another aspect of the much-debated fine-tuning problem in physics (which we already encountered in chapters III.2+3).

The inflationary model lends itself to further speculation. For instance, one might also assume that not just one initial false vacuum existed, but many. Eventually, many universes could exist, each originating from different vacuum bubbles randomly created by quantum chance. A theory known as *'chaotic inflation'* was proposed by a Russian-American theoretical physicist at Stanford University, Andrei Linde. It is also called *'eternal inflation'* because the inflationary phase of the universe's expansion lasts forever. Inside these universes, some local region can eventually originate new bubbles which undergo a new inflation phase. In this view, each universe can reproduce itself and the overall branching 'multiverse' – as it is nowadays fashionable to label it, especially by the media – grows like an infinite self-reproducing fractal.

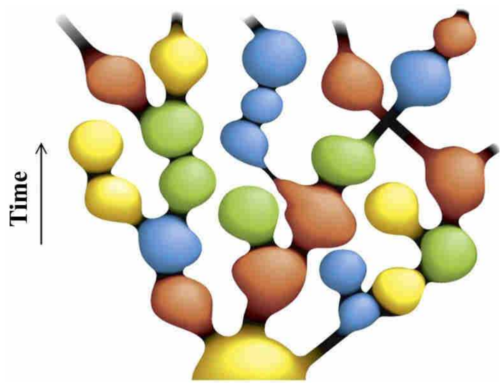


Fig. 83 Self-reproducing inflationary universe (A. Linde/Stanford University).

A very frightening doubt might arise at this point. If inflation theory is correct, what if our own universe, which is an almost entirely empty vacuum, is not really in a true but false vacuum state and enucleates a new vacuum bubble with a lower energy density? In principle, tiny bubble nucleation may occur everywhere and at any time. That is, a small patch of space in the universe can spontaneously decay due to quantum tunneling from its false to true vacuum state and, thereby, expand at the speed of light. If so, the inflating bubble would ignite a conversion into a lower state of energy in all the space regions it encounters on its path, finally devouring and destroying, without mercy, our entire universe. The life forms involved will not even have time to take notice, as there is no way to be warned beforehand. It is the ultimate doomsday scenario! Fortunately, it seems that the probability of such an event taking place is extremely low. It may be 10^{660}times the current age of the universe before such a catastrophic event will destroy us. This is much later than the lifetime of stars and galaxies. At that time, what will be left of the universe, if anything, will probably be only an empty and cold void at almost absolute zero temperature. Nobody would be left to worry about anything.

Because inflation cosmology is a theory which still needs final empiric confirmation, it is not exempt from criticism. The high-precision measurement of the CMB radiation by the European Space Agency 'Planck' satellite released in 2015 and, in the same year, the Background Imaging of Cosmic Extragalactic Polarization 2 (BICEP2) experiment at the South Pole seemed to confirm inflation theory. However, it soon became clear that the data could be explained without resorting to cosmic inflation. There is actually no unanimously accepted interpretation. To date, inflation theory remains an interesting scientific speculation but can still not be promoted to scientific truth. In 2017, a heated debate erupted between supporters of inflation theory and Paul Steinhardt (director of the Princeton Center for Theoretical Science), Anna Ijjas (Princeton Center for Theoretical Science), and Abraham Loeb (Harvard University's astronomy chair). The latter three published an article [28] that publicly cast doubts on what they called *"the long-cherished inflationary theory of the early cosmos"* and pleaded for new ideas. This soon received an angry reply from Guth, Linde, Hawking, and another 30 scientists, which resulted in a 'cosmic controversy' [29] that went viral on social media. This was probably because Steinhard is considered one of the main contributors to inflation theory, though he later turned his back on his own creation in favor of other cosmological models. It is a case which will certainly be of sociological interest to historians of science in the future.

At any rate, so far, inflation remains the favored quantum cosmology theory and the author refrains from repeating his personal stance on the ever-present discrepancy between modern theoretical physics and its apparent incapacity to go beyond speculations which lack empirical validation. It might be sufficient to compare this state of affairs of modern theoretical cosmology with that of the past mid-century. The rigor and crisp clarity of the last-century Big Bang nucleosynthesis theory, which leaves no space for wild speculations, along with the solid observational evidence supporting it, is something that inflation theory and the theoretical extensions of the SM to QG can only dream of. Clarity and certitudes are gone. Welcome to the quantum quagmire!

4. Quantum biology

Physicists and biologists did not believe (and most still don't) that quantum effects could have anything to do with biological processes and, as we shall see in the next chapter, with consciousness. After all, the quantum effects we observe are so tiny that it is hard to imagine how they could influence structures like a living cell, which is a macroscopic object compared to atoms and elementary particles. Any quantum phenomenon seems to be very unlikely to have any measurable effect and influence, just as one does not expect that a little feather hitting a truck will make the latter change its course. Moreover, if some quantum effect would arise (say, for instance, quantum superposition or entanglement phenomena among molecules in cells), the environment and the thermal heat bath to which every living creature is subjected would quickly cancel it by quantum decoherence. That is, Nature would confront the same problem we reviewed in the chapter on QC: External noise would immediately erase any quantum effect, rendering it useless.

However, Nature has surprised us so many times in the past and may do so this time, too. Quantum biology (QB) is a research field in its infancy but there are several indications that this may change in the not-too-distant future, though the speculations that QM may have something to say in biology are not new. Already in 1944, Schrödinger wrote a famous booklet entitled "What is Life?" in which similar conjectures were advanced for the first time. He asked if the ingredients that seem to make life so special and biology so different from chemistry might be traced to QP.

Also, the German physicist Pascual Jordan, who, among many others, contributed to the development of QM, and especially to the algebra of the commutation relations, speculated about the possible connections between quantum indeterminism and life. If the indeterminism of QP acting in cells could be amplified by some unknown mechanism at macroscopic scales, that could potentially explain the unpredictability of the complex biological structures that we are. Despite his achievements in theoretical physics, after WWII, Jordan's reputation became discredited because of his connections to Hitler's national socialistic ideology, as he had joined the Nazi party and its paramilitary unit. Probably because of this past political association, his works have been long dismissed and his conjectures about the possible connections between QP and biology have been forgotten.

In 1963, the Swedish physicist Per-Olov Löwdin suggested that the tunneling of protons between atomic nuclei might be a mechanism that could induce DNA mutations, and introduced the term *'quantum biology'*. This hypothesis must still be proven but it sounds plausible to conjecture that quantum tunneling is at work between nucleotides, i.e., the building blocks that make up DNA and that carry genetic information. The two DNA strands are held together by hydrogen bonds. In the process of replication, in which each strand makes a copy of itself, there is a small probability that a proton might 'jump' via a quantum tunneling process from one hydrogen atom to the other, thereby inducing a mutation.

In the 1980s-90s, British physicist Roger Penrose and the anaesthesiologist Stuart Hameroff advanced the hypothesis that QM could be at the base of the emergence of consciousness in the activity of neurons in our brains. Tiny microstructures in the cytoplasm of cells, called *'microtubules'*, are small enough to potentially support some quantum phenomenon. No evidence backs such a claim so far. We will return to this in the next chapter.

Not until 2014 did the British theoretical physicist Jim Al-Khalili publish a popular science book on QB with the suggestive title *"Life on the Edge: The Coming of Age of Quantum Biology"* [30], which made the subject known to the broad public.

All these attempts to look for a possible interface between the quantum world and biology rest on the fact that nowadays we know how molecular biology, genetics, and organic chemistry have been very successful in describing the processes of the functions of life in the microphysical domain. These are all about molecules essential to living organisms, many of which are made up of only a few atoms and stick together through chemical bonds which, ultimately, could be described only by resorting to the atomic quantum orbital models that are based on the laws of QM. (See also the brief account of the orbital concept at the end of the chapter on the state vector, Schrödinger's equation, and atomic physics in Vol. I.)

Our description here must necessarily be somewhat superficial, as a more detailed analysis would require a longer introduction to biology (also because the author is not a biologist and refrains from going too deep into the subject).

We will only hint at some of the recent lines of research, thereby hopefully encouraging the reader to go beyond this anecdotal introduction.

In fact, in the last decade, several instances of (though not completely conclusive) experimental evidence have emerged to indicate that QM might, indeed, play a role in biochemistry. For example, there is nowadays some consensus that enzymatic activity can't be explained without resorting to QP. Enzymes are proteins that act as catalysts in chemical reactions, accelerating biological processes in cells. Without enzymes, almost no living being could survive; they are among the most fundamental ingredients for life, at least as we know it. Some enzymes' chemical kinetics are faster than what classical chemistry can explain. Studies show how electron transfer (redox reaction) over long distances from one molecule to another (redox centers) could be achieved via quantum tunneling effects. This quantum mechanism plays an important role in enzymatic activity.

One of the strongest indications that quantum effects are at work in cells is photosynthesis. As everyone learns in school, photosynthesis is the chemical reaction in plants that converts the Sun's light into chemical energy. When photons hit the green-pigmented chlorophyll molecules, these induce, through an electron excitation mechanism, a charge separation triggering a complex chain of chemical reactions that convert them into chemical energy and transform carbon dioxide and water into glucose and oxygen. This energy transfer from the chlorophyll molecules to the cell's internal reaction sites has an efficiency of almost 100%, which can hardly be explained by classical chemical or physical principles. It looks like photosynthesis resorts to QM in order to achieve such efficiency. In fact, some studies suggest that long-lived quantum coherence combined with quantum tunneling may furnish the needed mechanism. Only by this mechanism can the packets of energy be transported to the cell by simultaneously following all possible paths to the reaction center.

Currently, research is seeking to determine whether QM could explain olfaction – that is, the biochemical reactions we use for the sense of smell. We know that olfactory receptors in the nose are quite efficient at binding odorant molecules. However, it is not entirely clear how the receptor molecules bind and recognize the different chemical substances. One theory states that the shape and size of the molecules are what determine the types of smells. These molecules are detected by the olfactory receptors onto which they must fit according to their shape and, by special signaling to the brain, decode it as different smells. In other words, according to this hypothesis, the recognition of smells is related to the form of the molecules carrying that smell. However, more recent tests suggest that molecules with very different shapes can lead to the same smell. Therefore, a new theory suggests that olfactory experience is not mediated by the shape of molecules but, rather, by their frequency of mechanical vibration, which is triggered by the quantum tunneling of an electron between the olfactory molecule and specific amino acids within the olfactory receptor. The latter theory is still somewhat controversial but experiments suggest that both are not mutually exclusive: The chemical composition, shape, and size of the molecules determine whether they are 'locked' into the receptors; their vibrational frequency then determines what we perceive as smell.

Another domain in which, quite surprisingly, QB might play an important role, and which has recently attracted attention, is migratory birds. It has been shown that at least some species are sensitive to the Earth's magnetic field – an ability called *'magnetoreception'*. The most striking example is the European Robin, which migrates in the winter from northern Europe to the Mediterranean Sea, and then back in the summer. For a long time, scientists wondered what the reference system of these birds could be such that they always find their way along their thousand-mile journey. The Earth's magnetic field seems to be the answer. However, it is a mystery as to how they can perceive such a weak magnetic field. Research on the European Robin suggests that the directional information about the Earth's magnetic field lines (more precisely, its inclination with respect to the surface) might be conveyed to the bird's brain by a mechanism relying on QM – in particular, on quantum entanglement. It is triggered by the light shining on the bird's retina, where magnetically sensitive protein radical-pairs molecules, called *'cryptochromes'*, reside.

Fig. 84 The European Robin

A radical-pair is made of two molecules with an odd number of unpaired electrons (recall how the Pauli exclusion principle determines the electronic structure in atoms) due to an ionization process photo-induced by light radiation. These two electrons are separated because each resides on its respective molecule but they are an entangled singlet. Of course, each of these electrons has a spin, which gives magnetic momentum to each separate molecule. And, just as we have seen that the electron's dynamic can be influenced by its own magnetic moment in an SG-experiment, here, also, the Earth's magnetic field can change the spin state. Therefore, the radical-pair mechanism explains magnetoreception insofar as a magnetic field can affect the radical-pair's chemical reactivity, which, in turn, determines the speed at which other chemical reactions and products are formed. In short: The direction of the Earth's magnetic field can, through electron entanglement, catalyze the biochemical reactions in the robin's eye. These reactions are then signaled to its brain. We will never know unless we become a European Robin but perhaps, thanks to QM, migrating birds might literally 'see' magnetic fields and orient themselves along these field lines.

QB is far from being an established and accepted science. So far, the evidence supporting it has gone beyond mere speculation but it isn't conclusive. Many physicists and biologists remain skeptical. As with QC, QB isn't really a new field of research. However, the decisive difference between the two research lines is that, while QC has been intensively pursued and heavily funded by billions in investment in the last couple of decades due to its potential practical and commercial applications, QB remains a largely ignored science. It has never received serious attention apart from that of a few scientists and a handful of research institutions. Thus, it has remained very limited in terms of scope, funds, and time. One of the reasons for this is that nobody has been able to come up with a convincing mechanism that could prevent quantum decoherence due to thermal noise and environmental interference. On the other hand, even if one day we discover quantum phenomena influencing or regulating life functions, it would not be clear what this knowledge could be good for. We live in a society driven by pragmatism and utilitarianism. If a relatively small field of research, however interesting and original it might be, isn't seen as useful for some practical and commercial application, it rarely gets funded. Pure science for the sake of knowledge is out of fashion. One can only hope that QB will receive more attention in the coming years. It would be interesting to know for sure whether QM does, or does not, play a role in the development of life. If it does, this could open an entirely new line of research leading to a paradigm shift in biology and potentially also to applications. The history of science shows that pure science can – and, indeed, in most cases does – lead to unexpected applications. Galileo, Newton, Einstein, and Planck, like many other great physicists, were driven primarily by a thirst for knowledge and not by the prospect of applying their discoveries. And yet, without their discoveries, we would still be stuck in the technological middle ages. Despite its inception half a century ago, QB remains a non-mainstream science. One can only hope that this will change soon.

5. Consciousness and quantum mechanics: myth or reality?

It was almost inevitable that the conjecture surrounding QB would sooner or later have been applied to the function of our brains. The human brain is a biological processing unit made of about 100 billion neurons. If QM plays a role in biological cells, and because neurons are nothing other than cells specialized for cognitive functions, it is hard to escape the temptation to extend the potential role of QB to neurons as well.

Moreover, a philosophical problem has plagued philosophers since the 17th century, the time when René Descartes stated, *"cogito, ergo sum"* – namely, a *'mind-body problem'* which ultimately boils down to the problem of consciousness. Nowadays, modern neuroscience dismisses Descartes' ideas about the mind and consciousness. Thinking, being, and perceiving are separate categories that can't be summarized by a slogan. However, what is particularly noteworthy is that the concept of *'consciousness'* continues to elude a definition and clear scientific categorization and explanation.

Despite the huge amount of data and discoveries of neurophysiology and neuropsychology, driven mainly by technological advances that allowed for the mapping and monitoring of the brain's activity, when it comes to the so-called *'hard problem of consciousness'*, one can safely say that not much progress has been made in the last four centuries. In modern philosophy, the debate over consciousness was summarized by the Australian philosopher David Chalmers, who distinguished between the easy and hard problems of consciousness. The relatively easy problem of consciousness is that which concerns the explanation of how the brain processes information and environmental stimuli or focuses attention. To specific cognitive functions, one correlates a specific neuronal activity in the brain. This is, so to speak, an 'easy' problem because the correlation has been extensively investigated using modern brain measuring and imaging techniques (e.g., electric and magnetic encephalography, neuroimaging with different computer tomography and magnetic resonance methods, etc.).

However, the hard problem has more of a philosophical nature and relates to the question of why a biochemical activity, however complex it may be, should give rise to an experiencing subject that perceives qualities. It looks like a complicated interaction of neurons does give rise to an individual that seems to be a separate "I", having subjective sensorial perceptions of pain and pleasure and registering colours, smells, tastes, touch, and sounds, which philosophers call *'qualias'*. The information from the environment is not experienced just as a stream of data of abstract bits, as it is encoded in our neural networks, but also as a qualitative phenomenal lived experience. Also, we do not perceive our mental activity as a program running on a set of logical circuits but, rather, as a content of insights that can span a spectrum beyond analytic representations of images and feelings and that find no parallel in a computer program or any algorithmic structure. Strictly speaking, a materialistic physicalist science should completely reject these qualias as pseudo-scientific woo-woo and give them no more credence than the existence of unicorns or goblins. This would certainly be the case if not for the fact that every scientist can verify, by him/herself, that she/he is more than a data-crunching robotic zombie. For some unknown reason, we are living organisms that have subjective experiences. If modern science accepts the existence of consciousness, as characterized above by the ability to perceive qualias, it is not because of any empiric external evidence, which is instead completely missing, but only because of an 'inter-subjective consensus'.

Moreover, there are also some peculiar aspects that characterize consciousness. For example, it behaves as something which is quite indeterministic and unpredictable. Is our free will real or only an illusion that arises from a complicated biological machine? Or, to put it in the quantum parlance: Are there hidden variables in our brain processes that determine our behavior according to a Laplacian view, or are there none and are we really just as free as an electron in the double slit experiment is 'free' to choose its interference fringe? The superdeterminists such as Bell or the supporters of BM would probably tend to believe that there is no free will because, in their interpretation of QM, even the electron's behavior is predetermined by hidden variables. Others, who accept an indeterministic and non-reductionist view of reality, might be more in favor of the existence of free will. In this latter case, if QM is a truly indeterministic theory without hidden variables, it would naturally accommodate free will and explain the unpredictability of consciousness beyond a mere manifestation of complexity.

Further, nobody knows how the brain manages to bind its perceptions of thoughts, feelings, and the environment into a unique and undivided whole to which it associates a meaning. This is another problem, called the *'binding problem'*. For instance, when we look at an image, we do not consciously register each pixel one by one and begin to make calculations or ponder on it separately. Instead, we just see, almost instantly, an image to which we associate a unified meaning, being itself a conglomerate of meanings at once. For example, when we see the image of a giraffe, we know at once that it is an animal (a categorial meaning) with a long neck (a morphological meaning) and a dark-yellow leopard-like colouring (a chromatic experience). Even if not displayed in the image, we eventually append to it (more or less unconsciously) the environment in which it lives, such as the savanna or a zoo (a mental figure). This is more than the sum of the parts and much more than pattern recognition (something artificial neural networks can already do). Instead, it is about the unified lived experience of a subject that, from a batch of seemingly disordered pixels and other data, suddenly 'collapses' it in the perception of the emergence of a meaning which, however, inherently retains a holistic cognition. The fact that meaning is not inherent in the objects 'out there' but must emerge due to the binding skill of our consciousness, is easily recognized through the famous Gestalt figures, in which different meanings could be 'entangled', such as those in which we see, at first glance, two faces and then, shortly thereafter, a vase. Nobody knows how the brain manages to mimic this unity of consciousness. Despite the extraordinary progress of neurobiology and brain imaging technologies, in this respect, almost no progress has been made.

The fact that these mysterious holistic aspects of our cognition are reminiscent of quantum effects has led some to conjecture that perhaps quantum coherence, with its indeterministic and non-local aspects, might be secretly at work in our brains. Maybe our brains are, at least partially, like a quantum computer able to process some information all at once. The subjective experience may be related to the complex dynamics of quantum neural networks.

The first conjectures in this direction came in 1961. Eugene Wigner (1992-1995), a Hungarian-American physicist, proposed an extended form of the Schrödinger's cat paradox, called the *'Wigner's friend paradox'*. In a thought experiment, Wigner imagined how a physicist in a laboratory (say, Wigner himself) does not directly make the observation of a quantum system in superposition (say, a particle in spin superposition or a cat which is dead or alive) but, rather, lets another physicist, his friend, perform it. Later, this friend will communicate the result to him. In the context of Schrödinger's cat, one can place Wigner's friend inside the box with the cat, the radioactive material, the Geiger counter, and, of course, the poisonous ampoule. When Wigner's friend takes the measurement, we know that state reduction occurs, and one of the two possibilities will be realized (spin-up or spin-down, or the dead or alive cat). However, according to the rules of QM, seen from the outside, before the friend's measurement, Wigner had to regard his friend in a state of superposition as well – namely, the friend who has information corresponding to one or the other outcome simultaneously. That is, Wigner had to consider the whole laboratory as a joint quantum system, the physical object plus the friend. Taking the example of Schrödinger's cat, this means that, for Wigner, the box contained an alive cat and his friend who had measured it being alive in superposition with a dead cat and the friend having measured the cat being dead. If, now, Wigner asked his friend about the state of affairs, he would get either the former or latter case as an answer. However, this entails that it only at the time of Wigner's question did the state collapse take place. According to the rules of QM, as long as no measurement takes place, the system must still be considered as being in a quantum superposition of the alternative outcomes. For Wigner's friend, the quantum collapse of the wavefunction occurred much earlier than what Wigner observed. There is, therefore, a discrepancy between the temporal sequence of the state reduction according to the two observers. Because there is no known physical law or rule that lifts Wigner or Wigner's friend to the status of a privileged observer, a paradox arises. As you can see, this is an extension of, or another perspective on, the measurement problem. It raises, again, the questions of when exactly the collapse occurs and whether it is actually a real physical process or only an abstract mathematical description of reality.

On several occasions, the author refrained from introducing the notion that the observer could have any supposed role in QM. First, QM is at work throughout the universe without the need for any observer. Second, if so, things could not be as simple as that, in the way that the popular media like to present it. At any rate, given the aforementioned paradoxical situation, Wigner pointed out that a way out of the impasse could be to posit that the physical act of measurement isn't sufficient to cause a quantum collapse; rather, the conscious awareness of an observer is necessary. As long as no conscious observers are involved (say, only microscopic quantum particles instead of cats, friends, and other physicists), they are effectively in a permanent superposition state all along the temporal sequence of the process. If, instead, a conscious observer is involved, that would instantly cause the collapse of the wavefunction. If Wigner's friend were conscious, the question Wigner posed was only just a second measurement on a system that had already projected onto its eigenbase. According to Wigner, *"the being with a consciousness must have a different role in quantum mechanics than the inanimate measuring device"*.

The Wigner's friend paradox raised the question of whether QM may have something to do with consciousness. Might consciousness be even more fundamental than quantum phenomena? Maybe quantum collapse is an act of consciousness itself. (This latter view is the author's preferred 'interpretation' of QM and will be taken up in a separate volume of a more philosophical character.)

This idea stayed in the background for several years but later paved the way to further speculations about the possible connections between QM and conscious experiences. Henry Stapp, an American mathematical physicist, argued that consciousness might precede matter in being fundamental to the universe. Our mental processes are, in themselves, due to quantum collapse, while free will is the manifestation of the quantum mechanical effects in the brain. [31] He invoked the quantum Zeno effect (see Vol. I) as evidence that processes can be delayed or modulated in their temporal occurrence by a conscious act. Stapp's worldview is that of a panpsychist – namely, that consciousness and mind are a primordial property of matter.

An approach not too different from the wavefunction collapse was taken by the already-mentioned Roger Penrose and Stuart Hameroff [32]. They argued that consciousness and mind cannot arise due only to an algorithmic computation. The brain's function could be non-computational and governed at the microscopic scale by quantum effects inside microtubules. Microtubules are tubular-shaped macromolecules in the cytoplasm of cells providing for their structure and shape. The inner diameter of a microtubule is about 12 nm (12×10^{-9}m) – that is, the size of the order of a couple of Caesium atoms. At this scale, quantum effects could become relevant and, if quantum coherence could be sustained for a sufficiently long time, eventually contribute significantly to neural activity. In addition to the electro-chemical signaling, Nature might have found, by this, a way to keep complex micro-systems in an entangled state and enact some form of quantum computing information processes inside the brain, conveying to our conscious experience and cognition its characteristic aspect that we know. Of course, in Penrose's and Hameroff's model, the problem of the environmental-induced decoherence remains unsolved.

More recently, some [33] [34] have proposed that a mechanism similar to that of photosynthesis or magnetoreception might occur in and between neurons. There is convincing evidence in support of the fact that some chemical reactions inside cells, with neurons being no exception, are capable of emitting *'biophotons'* (with the mitochondria being likely sources). We know that the neuron's activity is mainly of an electric and biochemical nature but light emission might also play a role. We know that light-sensitive proteins exist in the brain and could potentially function as single-photon detectors. Where photons are emitted or absorbed inside living tissue, the same or a similar QB of photosynthesis and magnetoreception might be at work. For instance, some imagine a neural complex entangled state of spins generated through the exchange of photons. Entangled spins and photons are somewhat less susceptible to the thermal decoherence effects which could potentially be a partial solution to the problem posed by the ever-present environmental quantum noise. If Nature is able to provide a mechanism by which the coherence can be preserved for time ranges of a subjective experience (milliseconds to a second), that could potentially furnish an indication.

These were only some of the many speculations surrounding the hypothesis concerning the possible connections between QP and consciousness. All these hypotheses raise many interesting questions in neuroscience, biology, biophysics, and philosophy. However, these are even further than QB from being accepted and established scientific facts. It is not at all clear whether QP and consciousness have something to do with each other, and several scientists dismiss this hypothesis altogether. Again, it is difficult to imagine how that warm grey matter could contain and sustain, even in principle, long-lived phenomena of that extremely fragile phenomenon that is quantum coherence. Few believe that decoherence at room temperatures in such a mushy macroscopic object could be prevented from settling in almost instantly. Moreover, most do not feel that it is necessary to invoke QM in the brain because they are interested only in solving the easy problem of consciousness. There is otherwise no reason to believe that any other cognitive function could be explained simply by the classical laws of physics.

At any rate, what could the practical applications be? Neurobiologists, psychologists, or physicians are not interested in philosophical ruminations; they must heal their patients. The pharmaceutical industry looks for chemists and biologists who manufacture new drugs to sell on the market; it doesn't look for philosophers or physicists who speculate about the nature of consciousness. And, of course, nobody is willing to finance such a line of research. Governments, academic institutions, and foundations care about quantum consciousness even less than they do about QB.

While in recent years, the research on QC and the possible existence of quantum effects in biological microstructures has given a new impulse to this topic, we must be aware of the fact that the role of QP in the brain is more a topic about which popular magazines like to talk, rather than being a scientific, actively engaged, organized form of research. Some dismiss it altogether as pseudo-science. As we have already tried to point out, we must be able to distinguish between a hypothesis and an established fact. One must be cautious to discriminate and see where an experimental fact ends and where one's own speculation begins, being in particular very careful to not mix truth and reality with one's own preferences and ideological stance. This is especially important when we hope and desire that some particular model of the world is the 'truth' and 'reality' (whatever these words might mean). While being interesting and serious domains of research, QB and quantum consciousness are, nevertheless, all too frequently invoked as established facts. Too many use them to support their personal opinions and interpretations of the world. Frequently, they fall into pseudo-scientific claims. How this attitude has led to a plethora of pseudo-scientific theories and what we urgently must learn to be careful about will be the subject of the next concluding chapter.

6. Quantum woo and physicalism

Nowadays, the word 'quantum' is widely abused but very much in fashion. People have invented 'quantum healing', 'quantum astrology', 'quantum architecture', 'quantum psychology', and even the 'quantum Jesus' and any sort of 'quantum nonsense'. All this has flooded the media and attracts lots of interested readers from all over the world. It seems that everything prefixed with the label 'quantum' automatically gains authority and becomes a guarantee and seal for the ultimate scientific truth. Nothing could be further from the truth.

Unfortunately, most of these new-age pseudo-scientific ruminations seem to have convinced the vast majority of non-physicists that, indeed, QP has definitely proven silly claims such as the existence of the soul, that one's health depends on the laws of QM, and, as everyone knows all too well, that QP shows how the outcome of the experiment depends on the 'role of the observer'. It is also well-known that in a networked society, if one repeats a falsehood a sufficient number of times, it will quickly spread throughout the collective and gain the status of accepted truth. Especially in the information age of social media, these claims can make it through more easily because they profit from the ability to use powerful technological means to cause ignorance to become viral, while the lack of discrimination among readers takes care of the rest. This is one of the reasons why this book came into existence: to establish a firm intellectual basis for the foundation of QP for those among the masses who are not necessarily professional physicists but who nevertheless are willing to make an effort to go beyond fairytales. Only once the non-experts know what QP is really about can they safely distinguish and discriminate between a serious scientific claim and more or less pseudo-scientific fantasy.

For example, let us briefly take up the idea of a possible connection between QP and medicine. Most of the 'quantum healing' proponents resort to the recent discoveries and relatively novel line of research of QB. We have outlined how, though not every biologist or physicist resonates with it, QB and the connections between quantum phenomena and consciousness can nevertheless be considered a scientifically sound line of research or, at least, a serious philosophical inquiry. That QM might play a role in cellular processes is now a widely accepted possibility, and that deep down even our brains might depend on quantum phenomena is a conjecture that several scientists take seriously. However, so far, any further extrapolation and desire to jump to conclusions must be done with great care. Though science might one day prove that our microscopic cells are driven by entangled electrons or that neurons fire according to quantum principles, this would still be very far from showing that our health – that is, the physical state of our macroscopic bodies – can be described and explained with quantum effects and even less that any sort of 'quantum therapy' would deliver some miraculous cure. Who writes is not a physician but, as one doesn't need to be a rocket scientist to know that you can't get to the moon with a high-altitude balloon, similarly, you don't need a Nobel in medicine to understand that, at least according to present knowledge, any parallel between QP and healing is a dubious – or, at least, a much-too-farfetched – extrapolation. Even in the eventuality that this might one day change, one can't take seriously the building – upon such a scarce and weak connection – of an entirely new 'science' which completely lacks any experimental evidence or any other type of support.

Though there are some exceptions (most notably, physicist Fritjof Capra, the best-selling author of the book *The Tao of Physics*, and who indeed offers some interesting philosophical points), most authors who write books about the supposed connections between QP and some sort of metaphysical effects in our daily lives are not physicists and don't have any training in the field. Most of these self-improvised scientists simply use wordplay and appealing metaphors to jump to conclusions that are not supported by logical and empiric evidence. The emblematic case is that of non-locality. Because QP is a non-local theory, this is immediately taken as final proof that the superluminal transmission of information is possible. Mixing this with unverified claims about QB and quantum consciousness, one jumps to the conclusion that telepathy, if it exists, is a superluminal quantum effect of instant transmission between mind and mind. However, this ignores the fact that physics has already shown this to be impossible. As you might recall, we explained it as being a gross misunderstanding; nothing in QP allows for FTL transmission of information (as illustrated in the chapter on FTL transmission in Vol. I).

Then, the fact that QP uses terms such as 'quantum fields' and 'energy field' together with its non-local holistic aspect of quantum entanglement is, for many, more than sufficient to extrapolate to the famous 'law of attraction' – the theory that says positive thinking will attract positive events to our lives (and that negative thinking will attract negative events). In particular, this latter 'quantum mysticism' was propagated by very successful documentary films like *What the Bleep Do We Know?* and *The Secret*, which had a worldwide influence and were followed by a wealthy business model with a plethora of books, conferences, and online and offline courses about how to become rich and happy by applying the law of attraction. Personally, from a purely psychological perspective, I'm inclined to also believe that positive thinking can help us have a better life. However, there is no need to invoke QP; it can be tested much better by experience rather than by resorting to particles and wavefunctions.

This, like many other quantum woo, owes its success to a weakness in critical thinking and discernment that plagues our post-truth society. These pseudo-scientific theories exploit the inability of the vast majority of the audience to distinguish between a verified scientific fact, a sensible theory, a good hypothesis, a simple conjecture, an unwarranted extrapolation, and outright fantasies. All these are shades and gradations of truths, partial truths, or untruths that most are unable to distinguish and separate from each other. Many throw everything (QP, medicine, biology, psychology, astrology, etc.) together into the same pot because so many seem to believe that this is a more 'holistic', 'democratic', and 'open-minded' approach which brings us closer to the truth. It is a widespread philosophical relativism that is very much in fashion. It rests on the assumption that there are no truths or falsehoods, that there is no good or bad, only equally valid 'points of view'. Whatever one says, even if it is based on firm scientific evidence, it is just an 'opinion' that is no more and no less 'true' than any other viewpoint. Saying that the Earth is flat is acceptable; it is an understanding of reality just like any other that we should democratically embrace. After all, the wave-particle duality has shown how everything can be both true or untrue, or that quantum entanglement shows that 'we are all one' and, therefore, must love each other. And the discovery of the zero-point energy is proof of the existence of a worldwide conspiracy of physicists to hide an inexhaustible source of energy out of nothing and from which humanity could potentially benefit. Right?

Whenever you read titles which begin with phrases such as “science got it all wrong” or “scientists finally admit” or, again, "final proof of the existence of" something that science has obviously kept secret or denied for a long time and that now has been exposed, then you must beware and raise your antennas. This is especially true if the author makes such a claim in the title of the article, which he/she refuses to subject to a serious peer-reviewed journal. There are hundreds of self-declared ‘experts’ in the field, so-called ‘cranks’ who are welcomed by unserious publishers eager to sell their magazines or who crawl the Internet looking for a public that declares them the ‘newfound Einsteins’ that mainstream science ignores. There is an overabundance of crackpots who usually have never seriously studied physics and who have no clue what they are really talking about but who present their own ‘theories of everything’, claiming to have discovered the ultimate ‘world-formula’. Planck, Einstein, Schrödinger, Dirac, Pauli, and all the past and present scientists got it wrong: The world has to know another truth. One of their typical psychological strategies is that they don’t even try to use rational arguments to convince their potential audience; instead, they play the part of the victim, lamenting that their ground-breaking theory is not taken seriously because of an organized conspiracy of mainstream science. It is something that easily pays off in a post-truth era dominated by conspiracy theories, fake news, and social media allowing for viral ignorance.

This doesn't mean that the author negates the possibility of metaphysical phenomena such as telepathy, self-healing powers, or holistic causes and effects in life. Quite the contrary, he believes that they, indeed, exist and has even written extensively about self-healing. However, we cannot resort to QP in any of its present forms to prove it, much less explain it. Though the multiplication of crackpot theories is an interesting social phenomenon, we won’t go into this sort of controversy any further. If you have read this far and still are not convinced, we will certainly not be able to convince you otherwise. If, instead, you have the feeling that this course on QP did furnish you with some basics that allow you to distinguish between an irrational theory and physics, then the aim of this book has been realized and no further arguments pro or contra something are necessary.

Having said that, the author believes that (perhaps to your surprise and/or disappointment) this analysis can also be turned upside down. While it is true that a naive, unreflective, and almost superstitious attitude dominates large strata of a scientifically uneducated society, it is also true that there exists a (more or less undeclared) strictly materialist atheistic militant movement that bases its reasoning and conclusions on unaware assumptions, unwarranted logical inferences, and materialistic dogmas. This *'physicalism'* is an attitude which is qualitatively not much less unreflective, simplistic, and ultimately ignorant than the former.

One of the most common fallacies of the materialist thinker who drives to different forms of scientism is to assume (mostly without being aware of it) that the intellectual rational analysis of that little homo sapiens which has appeared on the evolutionary scene just 'yesterday' is the ultimate tool of knowledge that finally describes reality as it is. Paradoxically, this also becomes apparent among the 'spiritual quantum-community', in which the urge and desire to describe and reduce every mystical experience via a rational and material explanation always resurfaces. However, the mind is just one cognitive function that emerged from an evolutionary process which almost certainly did not reach its peak. Evolution has reached just one step in between the ladder of the development of life. Not only would it be presumptuous to believe that the human mind is the only and final instrument capable of knowing the truth of things but it would also be outright unscientific because that would contradict the principles of Darwinian evolution – just that theory to which scientific materialism clings in the first place. The idea that, apart from our rational and scientific mind, there could be something which goes beyond mind, comes not to their mind (no pun intended). Could there be something that goes beyond the emotional, the rational, the infra-rational, being supra-rational?

The preferred escapade of the physicalist's thought is to state that there is no need to invoke mystic or subtle phenomena beyond material existence. These are, so goes the argument, 'non-necessary hypotheses'. It is an argument based on the principle of Occam's razor, or the *'law of parsimony'*, which states that one should opt only for explanations with the fewest assumptions. The extra-physical assumption is considered unnecessary because science has explained so many things we previously thought could be elucidated only through metaphysical thought that there is no reason to believe it will not do so again in the future. It is only a matter of time. Maybe we still need another four centuries but science will do so for everything else which still has not found its accommodation inside a frame of a rationalistic and material thinking reason. However, apart from the fact that this attitude indirectly (and unconsciously) restates the priority of human reason as the ultimate tool of cognition against the Darwinian concept of evolution itself, it assumes that since the inception of science (say, about four centuries ago), it has steadily progressed towards an ultimate knowledge and truth of things, though without reaching it completely (one of the main ideas of the philosopher of science Karl Popper, who considered science to be a theory of growing *'truthlikeness'*).

However, closer inspection of the historical facts shows how this is true only for the description of physical processes and not for what these processes ultimately are. We know what the physical laws are (by the way, without being able to agree on how to define what a 'physical law' is) and what dynamical principles govern the evolution in space and time of these processes, but we have no clue what these processes are in themselves. In physics, we talk about particles moving in space-time subjected to some forces. However, categories such as 'particle' (which, as you well know, in QP is already a notion that becomes very fuzzy), 'space', 'time', and 'force' are not explained; they are objects set a priori and considered to be obvious and self-evident givens. We discussed how precisely this has now been recognized as one of the stumbling blocks to conceiving of a theory of QG. In the last four centuries, science has made no tangible progress in explaining the ultimate reality of these concepts which we regard as real and concrete.

We pointed out the difference between the easy and hard problems of consciousness. The former receives several plausible answers by means of the description of the processes, such as the neural correlates in our brains, and science (especially neuroscience) has indeed made great advances in the last three decades. However, the hard problem of consciousness has remained mysterious and impenetrable. In these regards, science did not progress, not even by an iota, in centuries.

After all, particular scientific and intellectual skills are not necessary to understand the difference between a process and its essence. Just think of a bunch of matter and ask yourself: What, ultimately, is the thing we call 'matter'? QP or atomic physics, and even chemistry, like any other science, does not answer – and doesn't even try to formulate – this question. They will tell you that matter is made of molecules, which are made of atoms – the atoms of elementary particles such as electrons, protons, and neutrons. The latter are composites of quarks and are held together by nuclear forces, and so on. However, this only shifts the question of the ultimate essence of things into another domain and never answers it. We imagine molecules, atoms, and elementary particles as a 'chunk of matter', only smaller. The notion of 'force' at the micro level, as the macro level, is the same: It is anything able to cause a change of momentum in time. However, nobody knows what, ultimately, a force is. When it comes to describing the essence of the processes and objects that physics deals with, it remains silent, as it has done for the four centuries after Galileo began to measure the motion of bodies on the inclined plane. Physics describes the process to which these 'things' are subjected and has nothing to say about their ultimate nature. And it never pretended to do otherwise because it unconsciously knows that it will never be able to do that.

There have been attempts to come up with this philosophical issue, most notoriously by Alfred North Whitehead, an English mathematician and philosopher of the first half of the 20th century, who tried to establish a *'process philosophy'* which posits processes, rather than matter, as fundamental. In his famous work, "Process and Reality", he sees the actual, existing world as a network of *'actual entities'* – that is, of events and processes as the ultimate and fundamental reality. What we perceive of this universe of actual entities are its temporally overlapping related *'atomic occasions of experiences'.* Whitehead, who was also well-versed in physics, tried his utmost to develop a philosophy by a strict, rational reasoning and a precise set of logical rules and expressions which got to the bottom of physical reality. However, ultimately, he could not get rid of metaphysical categories, as logical positivism couldn't, either. Quite the contrary, his philosophy seems almost to compel him to recognize that one can't evade the logical conclusion of the existence of an atemporal actual entity, which he couldn't refrain from calling 'God'.

Therefore, the human mind finds itself at a standstill. If it tries to look for the ultimate ontology of things without falling into metaphysics, it will have to admit that it can't go beyond a certain point – never, not even in principle. That's why when physicalists are confronted with questions about the essence of things, they play the Occam's razor card or move the 'no-need for unnecessary hypotheses' chess piece. Deep down, they know that otherwise they not only would have to recognize science's failure to get us even closer to the truth of these things but also that they would have to go within and look inside themselves. The lack of this need arises because it already posits, a priori, a lack of interest in answering these uncomfortable philosophical questions in the first place, which is an unconscious manifestation of the refusal to know themselves. Otherwise, the physicalist would not have failed to note how, in four centuries of scientific enquiry, there has been no progress in answering the fundamental philosophical questions about our existence, such as what is matter and force, space and time, consciousness and life, as these require a subjective investigation at a spiritual level. Once we realize this, the 'need' would become not only necessary but even unavoidable. The widespread quantum woo on one side and physicalism on the other side appear as the two complementary social phenomena of our times. The former is caused by a lack of intellectual discrimination which leads to gullibility, while the latter is caused by a spiritual blindness which has led to forms of materialistic dogmatism. They are just two sides of the same coin: the weakness of the human mind and spirit. This will be the leitmotiv of a coming publication of the author.

VI. Conclusion

This was our journey through the weird world of QP. We laid out the technical aspects, the theories, the experiments, and the philosophical speculations that surround modern QP in an attempt to furnish an overview that might not have answered every question but that, hopefully, left you with a deeper understanding than that which you had at the beginning. My sincere hope is that it furnished you with the promised introductory guidance in the essentials that everyone interested in the foundations of QP and its philosophical implications should have. It is an understanding that goes beyond what most of the hyped media do with distorted representations of QP that rely on disinformation and misrepresentation of facts or truths or on half-truths that can easily be sold as 'scientific facts' to an audience without any background. We hope that you are now in a position to appreciate what the real science is about and also to better consider some of the pseudo-intellectual representations of QP.

While this technical exposition in a two-volume series ends here, it merely laid the groundwork for the philosophical journey which is to begin. The author will publish another book that will investigate the connections between science, consciousness, and reality. It will investigate the limits of reason on the brink of a spiritual materialism.

The reason why, in the last chapter, we placed so much emphasis on the distorted approaches to QP is that modern spiritualistic interpretations of QP are based mostly on ignorance and misunderstanding, thereby completely missing the point. However, we also outlined the tendency of the materialist who is prone to similar misunderstandings, being unable to go beyond a strictly materialistic, rational approach to reality. It is believed that we must go beyond both worldviews.

What we tried to make clear was, on the one hand, what the down-to-earth, hard facts of science are, the ones that can't be dismissed, without indulging too much in speculations and phantasies that, ultimately, are only unaware manifestations of one's desire to find a theory that is merely intended to confirm one's own pre-established ideological and/or religious worldview. These are facts that any 'spiritualist interpretation' must take into account without falling into the crackpot category. On the other hand, we abandoned the reductionist, deterministic approach, which clings to any kind of local realism that most of the materialistic-minded scientists would like to recover. This is a tendency clearly evidenced in the plethora of interpretations of QM which, more or less implicitly, hope to resurrect some form of revised and amended Newtonian mechanics but which turned out to be not much more than a sterile intellectual exercise. We believe this to be an expression of the human mind's inability to look inward and surge to higher intuitive cognitions than what a limited mechanical and physical mind can do.

Now that we have set—through these two volumes—a basis that hopefully has clarified what QP is about, but especially what it is not, the interested reader can follow me in raising the levels to a more philosophical and, at times, a spiritual and teleological (not theological), approach. It is time, then, to look beyond both approaches and search for a 'third position'. This will be the topic of my next book.

VII. Appendix

A I. Mathematical Appendix

a. The sine and cosine function with Euler's formula

Recall Euler's formula defined on the imaginary unit circle (see also the appendix of the previous volume):

$$e^{i\theta} = cos(\theta) + i\,sin(\theta); \quad \textit{Eq. 43}$$

The sine and cosine functions can be defined on the complex plane as:

$$cos(\theta) = \frac{e^{i\theta} + e^{-i\theta}}{2}; \quad \textit{Eq. 44}$$

$$sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i}. \quad \textit{Eq. 45}$$

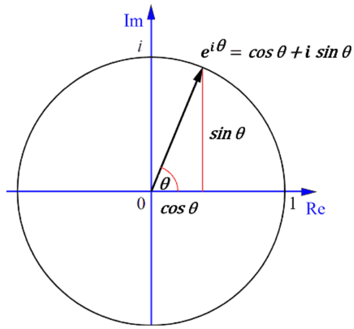


Fig. 85 The imaginary unit circle.

This is easy to check if you insert Eq. 44 and Eq. 45 back into Eq. 43.

b. The logarithm function

If you know what an exponential is, such as 3^7 ('three raised to the power of seven') or e^2 ('Euler's number squared'), it should not be too difficult to understand what its inverse is – namely, the logarithm of 3^7 being 7 in base 3 and the logarithm of e^2 being e in base e. In formula: $\log_3 3^7 = \log_3 2187 = 7$ or $\log_e e^2 = \log_e 7.3891... = 2$. The most common base for an exponent is, of course, base 10. It is only a historical convention which, however, relies on our biology; because humans have ten fingers, they use a ten-digit numeral system (0, 1, …9) to represent quantities. Most of our numberings and metric systems are based on base 10. Say the length measured in kilometers, divided in meters, centimeters, millimeters, that is, 1 km $=10^0$km = 10^3m = 10^5cm = 10^6mm, with the first term because any number a raised to the power zero is unity, always: $a^0 = 1$. The respective logarithms (in base 10) are $\log_{10} 1 = 0$, $\log_{10} 10^3 = \log_{10} 1000 = 3$, $\log_{10} 10^5 = \log_{10} 100000 = 5$, $\log_{10} 10^6 = \log_{10} 1000000 = 6$. These were easy examples but to what power must one raise 10 to obtain 16? The answer gives the logarithm (through your pocket calculator) as $\log_{10} 16 = 1.20411983...$, because $10^{1.20411983} = 16$. The human base ten is, however, not always the best with which to measure physical or mathematical quantities. Indeed, Nature seems to have some predilection for Euler's number base e=2.718281828… This is related to the fact that we live in a 'wavy universe', be it waves propagating in matter or quantum wavefunctions which can be described by sine and cosine functions with Euler's exponentials, as illustrated in Appendix A Ia. Somewhat easier to understand are the base two numbers which, apart from being presumably the preferred base of an extraterrestrial civilization with only two fingers, are the most natural choice for measuring the binary information with which all computers work. A byte is a register made of eight bits and can attain $2^8 = 256$ distinguishable physical states (see chapter IV.5), a 16-bit pixel can display $2^{16} = 65536$ colours, and a microprocessor working with 64-bit architecture can address ca. 1.8×10^{19} memory cells (that is, 18 quintillions of values). Of course, $\log_2 256 = 8$, $\log_2 65536 = 16$, and $\log_2 1.8 \times 10^{19} = 64$. Note how, in particular in the latter case, the logarithm functions serve well in reducing, to a humanly manageable order of magnitude, something which otherwise would result in extremely huge numbers. It is therefore also useful to display graphs which involve very small and very large numbers at the same time with one or both axes in logarithmic scales. (See, as an example, Fig. 79.)

Having clarified this, we can proceed with a bit more rigorous definition. The logarithm function is the inverse function of the exponential. That is, given a number x, the logarithm is the number a to which the base b must be raised to produce x: $x = b^a \rightarrow log_b\, x = a$.

On logarithms, tomes have been written since J. Napier introduced it in the 17th century, and we won't go into further details than this intuitive introduction provides. Only a couple of properties of the logarithm function should be mentioned and they turn out to be useful in the present book.

First, the logarithm is the only additive function that exists – namely, the logarithm of the sum of two numbers a and b is the product of the logarithm of each number as:

$$\log a \cdot \log b = \log(a + b).$$

This is an important aspect when it comes to measuring the entropy of a system which is supposed to be an additive quantity ('the whole is the sum of the parts'), something which no longer holds in QP

Then, the logarithm of an inverse power n is the negative logarithm of n, that is:

$$\log\frac{1}{n} = \log n^{-1} = -log\, n.$$

This should explain our observation for the equally probable values in the Boltzmann-Gibbs entropy Eq. 24 reducing to Boltzman entropy Eq. 23. If $p_i = \frac{1}{W}$, then: $\sum_i^W p_i \cdot log\, p_i = \sum_i^W \frac{1}{W} \cdot log\, \frac{1}{W} = -\sum_i^W \frac{1}{W} \cdot log\, W = -\frac{W}{W} \cdot log\, W = log\, W$.

c. Matrixes

What are matrixes? To put it bluntly, they are not much more than two-dimensional collections of numbers. Scalars are just single numbers (for, example a temperature value)—that is, a 0-dimensional array. Vectors (for example, an arrow with a length pointing towards a direction, like a velocity or an acceleration) are a one-dimensional array of numbers. A matrix is a two-dimensional rectangular array of numbers which are ordered in rows x columns and which can be useful in describing some physical quantities—for example, a scalar field like that illustrated in Fig. 34 left or, as we will see in the next appendix subsection, density matrixes representing an extension of the state vector in quantum mechanics to multiparticle systems. A general representation of a NxM matrix with coefficients a_{ij} (i=1…N; j=1...M) is:

$$M = \begin{pmatrix} a_{11} & \cdots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NM} \end{pmatrix}$$

For example, here are 2x3, 3x2, and square 2x2 matrixes:

$$M_1 = \begin{pmatrix} 1 & 7 & 5 \\ 4 & 9 & 13 \end{pmatrix}; \; M_2 = \begin{pmatrix} 2 & 6 \\ 3 & 22 \\ 19 & 8 \end{pmatrix}; \; M_3 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Special matrixes are the diagonal matrix (all zeros except on the diagonal) and the identity matrixes (a diagonal matrix with only ones on its diagonal), such as in the 3x3 square matrix case:

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 9 \end{pmatrix}; \; I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Symmetric matrixes are square matrixes which have the same coefficients above or under the diagonal, such as the 3x3 symmetric matrix:

$$S = \begin{pmatrix} 1 & 12 & 27 \\ 12 & 5 & 3 \\ 27 & 3 & 9 \end{pmatrix}.$$

The trace (Tr) of a matrix A is the sum over all its diagonal coefficients, that is: $Tr\, A = \sum_{i=1}^{N} a_{ii}$, for example:

$$Tr \begin{pmatrix} 1 & 5 & 7 \\ 9 & 2 & 6 \\ 1 & 2 & 3 \end{pmatrix} = 1 + 2 + 3 = 6.$$

Matrixes can be transformed and represented in different coordinate bases. They can be added or multiplied with each other or by a scalar. They can act on vectors, rotate it, or extend it. It is in matrix algebra that the concepts of eigenvectors and eigenvalues were first defined. The manipulation of matrixes follows strict algebraic rules and is a huge topic on which people have written tomes. For our purposes, however, these few lines to give you an intuitive understanding are sufficient.

d. Density matrixes

Generally, given a mixture of N pure states represented by their state vectors and respective probabilities $\{ (|\Psi_1\rangle, p_1), (|\Psi_2\rangle, p_2), \ldots, (|\Psi_N\rangle, p_N) \}$, the density operator is defined by the sum over matrices ρ_k as:

$$\rho = \sum_k p_k |\Psi_k\rangle\langle\Psi_k| = \sum_k p_k \, \rho_k(|\Psi_k\rangle) \,, \quad \textit{Eq. 46}$$

with $\rho_k(|\Psi_k\rangle)$ the k-th matrix with coefficients:

$$\rho_{ij}(|\Psi_k\rangle) = \langle e_i|\Psi_k\rangle\langle\Psi_k|e_j\rangle \,; \; (i, j = 1, 2, \ldots, D) \quad \textit{Eq. 47}$$
$$(k = 1, 2, \ldots, N)$$

The dimension D of the row x column matrixes is the dimension of the Hilbert space spanned by the eigenstate vectors $|e_i\rangle$ and where the state vectors $|\Psi_k\rangle$ are defined (or 'live').

To see what that means specifically and how one determines the density matrix, first recall how Dirac's bra-kets $\langle a|b\rangle$ are defined. These are a product of a complex conjugate row vector $\langle a|$ times a column vector $|b\rangle$ (see, again, the state vector and Schrödinger equation chapter in Vol. I). High school linear algebra tells us that if $\langle a|$ and $|b\rangle$ are orthogonal vectors (90° to each other), its Dirac bra-ket $\langle a|b\rangle = 0$. Fixing our ideas on the spin observable, the two spin eigenstates were represented by spinors (here, we omit any reference to a preferred direction x, y, or z):

$$|+\rangle = \begin{pmatrix}1\\0\end{pmatrix} \quad ; \quad |-\rangle = \begin{pmatrix}0\\1\end{pmatrix},$$

And, therefore, the four possible bra-kets' coefficients have values:

$$\langle +|+\rangle = 1, \; \langle +|-\rangle = 0, \; \langle -|+\rangle = 0, \; \langle +|+\rangle = 1. \quad \textit{Eq. 48}$$

<u>Example I</u>: Let us obtain the density matrix for the elementary case of a particle in a mixed spin-state (that is, *not* in state-superposition) represented by the statistical ensemble: $\{(|+\rangle, \frac{1}{2}), (|-\rangle, \frac{1}{2})\}$. The Hilbert space where our particle 'lives' is spanned (D=2→ i,j=1,2; N=2 → k=1,2) by the eigenstates $|e_1\rangle = |+\rangle$ and $|e_2\rangle = |-\rangle$ and we are considering it as having only two possible states, obviously defined by the very same state vectors $|\Psi_1\rangle = |+\rangle$ and $|\Psi_2\rangle = |-\rangle$. Then, from the right-hand side of Eq. 47 and using Eq. 48:

$$\rho_{11}(|+\rangle) = \langle e_1|\Psi_1\rangle\langle\Psi_1|e_1\rangle = \langle +|+\rangle\langle +|+\rangle = 1 \cdot 1 = 1 \,;$$

$$\rho_{12}(|+\rangle) = \langle e_1|\Psi_1\rangle\langle\Psi_1|e_2\rangle = \langle +|+\rangle\langle +|-\rangle = 1\cdot 0 = 0\,;$$
$$\rho_{21}(|+\rangle) = \langle e_2|\Psi_1\rangle\langle\Psi_1|e_1\rangle = \langle -|+\rangle\langle +|+\rangle = 0\cdot 1 = 0\,;$$
$$\rho_{22}(|+\rangle) = \langle e_2|\Psi_1\rangle\langle\Psi_1|e_2\rangle = \langle -|+\rangle\langle +|-\rangle = 0\cdot 0 = 0\,.$$

This gives us the first density matrix for state-vector $|+\rangle$:

$$\rho(|+\rangle) = |+\rangle\langle +| = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}.$$

With exactly the same approach (exercise!), one also obtains the density matrix for state vector $|-\rangle$:

$$\rho(|-\rangle) = |-\rangle\langle -| = \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}.$$

From this, it follows that the density matrix of Eq. 46 for the statistical ensemble $\{(|+\rangle,\frac{1}{2}),(|-\rangle,\frac{1}{2})\}$ is:

$$\begin{aligned}\rho(|+\rangle,|-\rangle) &= p_1\cdot\rho(|+\rangle) + p_2\cdot\rho(|-\rangle)\\ &= p_1\,|+\rangle\langle +| + p_2\,|-\rangle\langle -|\\ &= \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} + \frac{1}{2}\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix} = \begin{pmatrix}0.5 & 0\\ 0 & 0.5\end{pmatrix}.\end{aligned}$$

<u>Example II</u>: Now let us obtain the density matrix for the same particle of example I but in a spin-superposition state: $\{(\frac{|+\rangle \pm |-\rangle}{\sqrt{2}}, 1)$. The Hilbert space where our particle 'lives' is spanned again by the eigenstates $|e_1\rangle = |+\rangle$ and $|e_2\rangle = |-\rangle$ (D=2 → i,j=1,2) , but this time we are considering a single pure state (therefore, N=1→k=1) and the certainty given by probability p=1) and defined by only one (symmetric or anti-symmetric) state vector $|\Psi\rangle = \frac{|+\rangle \pm |-\rangle}{\sqrt{2}}$. Before plugging it all in Eq. 47, let us first evaluate the single terms:

$$\langle e_1|\Psi\rangle = \left\langle +\left|\frac{|+\rangle \pm |-\rangle}{\sqrt{2}}\right.\right\rangle = \frac{\langle +|+\rangle \pm \langle +|-\rangle}{\sqrt{2}} = \frac{1\pm 0}{\sqrt{2}} = \frac{1}{\sqrt{2}} = \langle\Psi|e_1\rangle\,;$$
$$\langle e_2|\Psi\rangle = \left\langle -\left|\frac{|+\rangle \pm |-\rangle}{\sqrt{2}}\right.\right\rangle = \frac{\langle -|+\rangle \pm \langle -|-\rangle}{\sqrt{2}} = \frac{0\pm 1}{\sqrt{2}} = \pm\frac{1}{\sqrt{2}} = \langle\Psi|e_2\rangle\,;$$

with the right-hand side equality because the commutation of Dirac bra-kets is nothing other than its complex conjugate (which, in this case, are just real numbers and therefore don't change), something that is also easy to check directly. Because N=1, there is only a single matrix in Eq. 47 whose coefficients are:

$$\rho_{11}(|\Psi\rangle) = \langle e_1|\Psi\rangle\langle\Psi|e_1\rangle = \frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}} = \frac{1}{2}\,;$$
$$\rho_{12}(|\Psi\rangle) = \langle e_1|\Psi\rangle\langle\Psi|e_2\rangle = \frac{1}{\sqrt{2}}\cdot\frac{\pm 1}{\sqrt{2}} = \pm\frac{1}{2}\,;$$

$$\rho_{21}(|\Psi\rangle) = \langle e_2|\Psi\rangle\langle\Psi|e_1\rangle = \frac{\pm 1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = \pm\frac{1}{2} \, ;$$
$$\rho_{22}(|\Psi\rangle) = \langle e_2|\Psi\rangle\langle\Psi|e_2\rangle = \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = \frac{1}{2}.$$

Therefore, the density matrix Eq. 46 is:

$$\rho\left(\frac{|+\rangle \pm |-\rangle}{\sqrt{2}}\right) = |\Psi\rangle\langle\Psi|$$
$$= \frac{1}{2}|+\rangle\langle+| \pm \frac{1}{2}|+\rangle\langle-| \pm \frac{1}{2}|-\rangle\langle+| + \frac{1}{2}|-\rangle\langle-|$$
$$= \begin{pmatrix} \frac{1}{2} & \pm\frac{1}{2} \\ \pm\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Example III: What about the density matrix of two entangled particles A and B? Remember, also, in this case, we are considering a single pure state vector (N=1, p=1), which you know well to be (for the type-II entangled photons):

$$|\Psi\rangle = \frac{|+\rangle_A |-\rangle_B \pm |-\rangle_A |+\rangle_B}{\sqrt{2}}.$$

The Hilbert space where the two particles 'live' is spanned by the eigenstates:

$$|e_1\rangle = |+\rangle_A|+\rangle_B = |+ +\rangle,$$
$$|e_2\rangle = |+\rangle_A|-\rangle_B = |+ -\rangle,$$
$$|e_3\rangle = |-\rangle_A|+\rangle_B = |- +\rangle,$$
$$|e_4\rangle = |-\rangle_A|-\rangle_B = |- -\rangle,$$

with the -kets on the right-hand side being a shorthand we will use interchangeably. Note that it is a four-dimensional abstract space (D=4; k=1) which means we have a single 4x4 matrix. Then:

$$\langle e_1|\Psi\rangle = \frac{\langle + +| \, (|+ -\rangle \pm |- +\rangle)}{\sqrt{2}} = \frac{\langle +|\langle ++\rangle|-\rangle \pm \langle +|\langle +-\rangle|+\rangle}{\sqrt{2}}$$
$$= \frac{1}{\sqrt{2}}(0 \pm 0) = 0 = \langle\Psi|e_1\rangle \, .$$

Following exactly the same algebraic approach one also obtains: $\langle e_2|\Psi\rangle = \langle\Psi|e_2\rangle = \pm\frac{1}{\sqrt{2}}$; $\langle e_3|\Psi\rangle = \langle\Psi|e_3\rangle = \frac{1}{\sqrt{2}}$; $\langle e_4|\Psi\rangle = \langle\Psi|e_4\rangle = 0$.

Then, the density matrix elements are:

$$\rho_{11}(|\Psi\rangle) = \langle e_1|\Psi\rangle\langle\Psi|e_1\rangle = 0 \, ; \quad \rho_{12}(|\Psi\rangle) = \langle e_1|\Psi\rangle\langle\Psi|e_2\rangle = 0 \, ;$$
$$\rho_{13}(|\Psi\rangle) = \langle e_1|\Psi\rangle\langle\Psi|e_3\rangle = 0 \, ; \quad \rho_{14}(|\Psi\rangle) = \langle e_1|\Psi\rangle\langle\Psi|e_4\rangle = 0 \, ;$$

$$\rho_{21}(|\Psi\rangle) = \langle e_2|\Psi\rangle\langle\Psi|e_1\rangle = 0\,; \quad \rho_{22}(|\Psi\rangle) = \langle e_2|\Psi\rangle\langle\Psi|e_2\rangle = \frac{1}{2}\,;$$
$$\rho_{23}(|\Psi\rangle) = \langle e_2|\Psi\rangle\langle\Psi|e_3\rangle = \pm\frac{1}{2}\,; \quad \rho_{24}(|\Psi\rangle) = \langle e_2|\Psi\rangle\langle\Psi|e_4\rangle = 0\,;$$
$$\rho_{31}(|\Psi\rangle) = \langle e_3|\Psi\rangle\langle\Psi|e_1\rangle = 0\,; \quad \rho_{32}(|\Psi\rangle) = \langle e_3|\Psi\rangle\langle\Psi|e_2\rangle = \pm\frac{1}{2}\,;$$
$$\rho_{33}(|\Psi\rangle) = \langle e_3|\Psi\rangle\langle\Psi|e_3\rangle = \frac{1}{2}\,; \quad \rho_{34}(|\Psi\rangle) = \langle e_3|\Psi\rangle\langle\Psi|e_4\rangle = 0\,;$$
$$\rho_{41}(|\Psi\rangle) = \langle e_4|\Psi\rangle\langle\Psi|e_1\rangle = 0\,; \quad \rho_{42}(|\Psi\rangle) = \langle e_4|\Psi\rangle\langle\Psi|e_2\rangle = 0\,;$$
$$\rho_{43}(|\Psi\rangle) = \langle e_4|\Psi\rangle\langle\Psi|e_3\rangle = 0\,; \quad \rho_{44}(|\Psi\rangle) = \langle e_4|\Psi\rangle\langle\Psi|e_4\rangle = 0\,;$$

This, finally, leads us to the density matrix of the two entangled particles:

$$\rho\left(\frac{|+\rangle_A|-\rangle_B \pm |-\rangle_A|+\rangle_B}{\sqrt{2}}\right) = |\Psi\rangle\langle\Psi|$$

$$= \frac{1}{2}|+-\rangle\langle+-| \pm \frac{1}{2}|+-\rangle\langle-+| \pm \frac{1}{2}|-+\rangle\langle+-| + \frac{1}{2}|-+\rangle\langle-+|$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \pm\frac{1}{2} & 0 \\ 0 & \pm\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

As an exercise, it is left to the reader to show that the type-I entangled particle has density matrix:

$$\rho\left(\frac{|+\rangle_A|+\rangle_B \pm |-\rangle_A|-\rangle_B}{\sqrt{2}}\right) = |\Psi\rangle\langle\Psi|$$

$$= \frac{1}{2}|++\rangle\langle++| \pm \frac{1}{2}|++\rangle\langle--| \pm \frac{1}{2}|--\rangle\langle++| \pm \frac{1}{2}|--\rangle\langle--|$$

$$= \begin{pmatrix} \frac{1}{2} & 0 & 0 & \pm\frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \pm\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

Regarding the physical interpretation of these matrixes, refer to chapter IV.6 on quantum information theory.

A II. Interference of light waves with different polarizations

The interference between two light beams having the same polarization (like in the case of the Young double-slit experiment) can be generalized to two waves with different linear polarization states. Recall how waves are described in time and space by their amplitudes and phases with the complex Euler numbers, as described in the appendix of Vol. I.

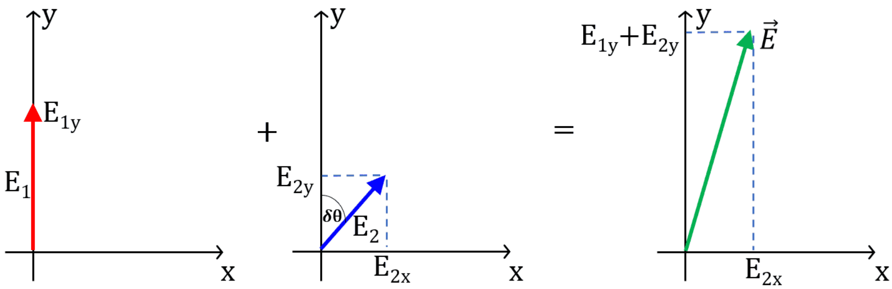


Fig. 86 Sum of two polarization vectors and its result.

Consider (Fig. 86 left) an EM wave with a vertical polarization vector—that is, with an electric field amplitude oscillating along the y-axis, $|E_1| = E_{1y}$, with angular frequency ω and none for the component E_{1x} along the x-axis:

$$E_{1x} = 0 \qquad ; \quad E_{1y} = |E_1|e^{i\omega t}$$

Similarly, consider the second vector in Fig. 86 as representing the polarization vector, another wave with an electric field amplitude $|E_2|$ having the horizontal and vertical components oscillating along both the x-axis and y-axis, E_{2x} and E_{2y} respectively, forming an angle $\delta\theta$ (with respect to the vertical vector E_1), with angular frequency ω and relative phase $\delta\phi + \delta\varepsilon$ (induced by the two slits path difference plus the eventual retarding optical plate phase shift):

$$E_{2x} = |E_2|\, e^{i(\omega t+\delta\phi+\delta\varepsilon)} sin(\theta) \quad ; \quad E_{2y} = |E_2| e^{i(\omega t+\delta\phi+\delta\varepsilon)} cos(\theta).$$

To simplify the notation, let us write, in more compact form, $\Delta = \delta\phi + \delta\varepsilon$. Then, the superposition of the two polarization vectors along the x- and y-axes is:

$$E_x = E_{1x} + E_{2x} = |E_2|e^{i(\omega t+\Delta)} sin(\delta\theta);$$

$$E_y = E_{1y} + E_{2y} = |E_1|e^{i\omega t} + |E_2|e^{i(\omega t+\Delta)} cos(\delta\theta)$$
$$= e^{i\omega t}(|E_1| + |E_2|\, e^{i\Delta} cos(\delta\theta)).$$

To obtain the contribution to the intensity of the EM beam from each component, we must modulus square it:

$$|E_x|^2 = E_x E_x^* = |E_2|^2\, sin^2(\delta\theta);$$

$$|E_y|^2 = E_y E_y^* = (|E_1| + |E_2|e^{i\Delta} cos(\delta\theta))(|E_1| + |E_2|e^{-i\Delta} cos(\delta\theta))$$
$$= |E_1|^2 + |E_1||E_2|e^{-i\Delta} cos(\delta\theta) + |E_1||E_2|e^{+i\Delta} cos(\delta\theta) + |E_2|^2 cos(\delta\theta)^2$$

$$= |E_1|^2 + |E_1||E_2|cos(\delta\theta)(e^{-i\Delta} + e^{+i\Delta}) + |E_2|^2 cos(\delta\theta)^2$$
$$= |E_1|^2 + |E_2|^2 cos(\delta\theta)^2 + 2|E_1||E_2|cos(\Delta)cos(\delta\theta),$$

with the last passage because of Euler's identity Eq. 44 for the cosine function.

Putting this together, to obtain the resulting vector, we can use the good old Pythagorean theorem. Noting that $cos^2\theta + sin^2\theta = 1$, it becomes:

$$|E|^2 = |E_x|^2 + |E_y|^2 = |E_1|^2 + |E_2|^2 + 2|E_1||E_2|cos(\Delta)cos(\delta\theta).$$

Recalling that the modulus squared of an electric field gives the intensity of a beam, we can label the intensity of the first and second beams as $|E_1|^2 = I_1$ and $|E_2|^2 = I_2$, respectively, to finally obtain:

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\, cos(\Delta)\, cos(\delta\theta). \quad \textit{Eq. 49}$$

If the two beams have the same intensity $I_1 = I_2 = I_0$ (say, because the two slits are equal), then *Eq. 49* simplifies to:

$$I = 2I_0 + 2I_0\, cos(\Delta)\, cos(\theta) \; = \; 2I_0[1 + cos(\Delta)\, cos(\delta\theta)]. \quad \textit{Eq. 50}$$

A III. Interference at detectors D_1 and D_2 in the delayed choice quantum eraser of Kim et al.

To see what really happens after the signal photon has been measured at detector D_0, we must continue to follow the idler photon along its journey towards detectors D_1 or D_2. We already discussed what is going on at detectors D_3 and D_4. Strictly speaking, the transmission of a photon across a beamsplitter induces a $\frac{\pi}{2}$ phase shift but because this happens at both beamsplitters BS_A and BS_B, we can ignore this phase shift. It affects the idler photon state function equally on both paths A and B, leaving unaffected the relative phase of the ket-vectors of $|\Psi_i\rangle_1$ or $|\Psi_i\rangle_2$ (Eq. 11 and Eq. 13). Therefore, from now on, we will only consider the idler photon after beamsplitters BS_A and BS_B.

Despite the apparently perfect physical path symmetry of the experimental setup of Fig. 22, one cannot say the same thing about the optical symmetry once the phase shifts induced by the reflection of the idler photon at mirrors M_A and M_B and beamsplitter BS are considered. A phase-shift anti-symmetry holds! It turns out that there is a small but decisive difference if one considers the photon's propagation along the optical path A reflected at mirror M_A and transmitted through beamsplitter BS towards detector D_1 or the same photon propagating along optical path B, reflected at mirror M_B and transmitted through the other side of the same beamsplitter BS towards detector D_2. So, let us follow the idler's propagation step by step.

First, keep in mind that when a photon (or an EM wave) is reflected by a mirror or beamsplitter, it undergoes a π-phase shift (180°-phase shift), while, as already mentioned, when it is transmitted through a beamsplitter, it undergoes a $\frac{\pi}{2}$-phase shift (90°-phase shift). In general mathematical terms, a phase shift φ applied to a state vector (or wavefunction) $|\Psi\rangle$ is represented by multiplying it by the complex Euler exponential $e^{i\varphi}$, that is, $|\Psi\rangle \rightarrow e^{i\varphi}|\Psi\rangle$.

With this, we can proceed by analyzing specifically how the photons of Eq. 11 or Eq. 13 behave. The expression of the photon's state in the rectilinear polarization basis $\mathcal{L} = \{|H\rangle, |V\rangle\}$ is not very interesting because we already know that these, being orthogonal, will not produce any interference pattern and reveal anything about the fringe and anti-fringe interference components. We can expect more insight by choosing to inspect the idler photon's state in the diagonal basis $\mathcal{D} = \{|45°\rangle, |-45°\rangle\}$.

Begin with the symmetric part of Eq. 11 or Eq. 13, the two possible idler photon states that result after collapse at detector D_0, that is:

$$|\Psi_i\rangle_1^{Sym} = |\Psi_i\rangle_2^{Sym} = \frac{|45°_i\rangle_A + |45°_i\rangle_B}{\sqrt{2}}. \quad \textit{Eq. 51}$$

Consider first the photon traveling along path A (after BS_A) reaching detector D_1, first by being reflected at mirror M_A (π-phase shift) and then being transmitted through beamsplitter BS ($\frac{\pi}{2}$-phase shift). Therefore, it undergoes a total $\frac{3}{2}\pi$-phase shift. Because $e^{i\frac{3}{2}\pi} = -i$, the idler photon's state vector on path A, $|45°_i\rangle_A$, transforms into: $|45°_i\rangle_A \rightarrow -i\,|45°_i\rangle_A$.

However, the idler photon is in superposition with itself and is also traveling along path B (after BS_B), reaching the same detector D_1 by being reflected twice, at mirror M_B (π-phase shift) and then at beamsplitter BS (π -phase shift). It undergoes a total 2π-phase shift. Because $e^{i2\pi} = 1$, the idler photon's state vector on path B, $|45°_i\rangle_B$, is invariant under such a transformation, that is: $|45°_i\rangle_B \rightarrow |45°_i\rangle_B$.

Therefore, detector D_1 will measure the quantum state:

$$|\Psi_{D_1}\rangle^1 = \frac{-i|45°_i\rangle_A + |45°_i\rangle_B}{\sqrt{2}}. \quad \textit{Eq. 52}$$

Now let's do this the other way around, towards detector D_2. The idler photon traveling along path A towards detector D_2 is reflected twice, at mirror M_A (π-phase shift) and then at beamsplitter BS (π-phase shift). It undergoes a total 2π-phase shift and, because $e^{i2\pi} = 1$, the idler photon's state vector coming from path A, $|45°_i\rangle_A$, is invariant under such a transformation, that is: $|45°_i\rangle_A \rightarrow |45°_i\rangle_A$.

However, being in superposition, the idler photon is also traveling along path B towards detector D_2 and is first reflected at mirror M_B (π-phase shift), then transmitted through beamsplitter BS ($\frac{\pi}{2}$-phase shift). It undergoes a total $\frac{3}{2}\pi$-phase shift and because $e^{i\frac{3}{2}\pi} = -i$, the idler photon's state vector on path B, $|45°_i\rangle_B$ is transformed into: $|45°_i\rangle_B \rightarrow -i\,|45°_i\rangle_B$.

Therefore, detector D_2 will measure the quantum state:

$$|\Psi_{D_2}\rangle^1 = \frac{|45°_i\rangle_A - i\,|45°_i\rangle_B}{\sqrt{2}}. \quad \textit{Eq. 53}$$

Eq. 52 and Eq. 53 are not the same. We obtained:

$$|\Psi_{D_2}\rangle^1 = -\,i\,|\Psi_{D_1}\rangle^1,$$

that is, they differ by a multiplicative factor, the negative imaginary number $-\,i = e^{i\frac{3}{2}\pi} = e^{-i\frac{\pi}{2}}$. (Recall how complex numbers are represented on the unitary complex circle; see also the mathematical appendix of Vol. I.) This can be interpreted as a relative phase shift of 90° of the wavefunction between detectors D_1 and D_2.

The question at this point is: What kind of signal will these photon quantum states produce at the two detectors? Fringes, anti-fringes, or a Gaussian profile? To see this, what remains to do is to modulus-square the two wavefunctions of Eq. 52 and Eq. 53. For the sake of brevity, let us simplify. First, omit the probability normalization coefficient which will play no role in the interference. Then, instead of Dirac notation, let us use the wavefunction notation and replace the kets of the diagonal polarization basis $\mathcal{D}$ as follows: $|+45°_i\rangle_A \rightarrow \mathcal{D}^+_{i,A}$, $|-45°_i\rangle_A \rightarrow \mathcal{D}^-_{i,A}$, $|+45°_i\rangle_B \rightarrow \mathcal{D}^+_{i,B}$, $|-45°_i\rangle_B \rightarrow \mathcal{D}^-_{i,B}$, and then also the final state functions at detectors D_1 and D_2 with $|\Psi_{D_1}\rangle^1 \rightarrow \Psi^1_{D_1}$, $|\Psi_{D_2}\rangle^1 \rightarrow \Psi^1_{D_2}$.

Then, Eq. 52 can be written as wavefunction:

$$\Psi^1_{D_1} = -i\,\mathcal{D}^+_{i,A} + \mathcal{D}^+_{i,B}.$$

Taking the modulus square:

$$\begin{aligned}
|\Psi^1_{D_1}|^2 = \Psi^1_{D_1}\cdot\Psi^{1*}_{D_1} &= (-i\,\mathcal{D}^+_{i,A} + \mathcal{D}^+_{i,B})\left(-i\mathcal{D}^+_{i,A} + \mathcal{D}^+_{i,B}\right)^* \\
&= (-i\,\mathcal{D}^+_{i,A} + \mathcal{D}^+_{i,B})\left(i\,\mathcal{D}^{+*}_{i,A} + \mathcal{D}^{+*}_{i,B}\right) \\
&= \left|\mathcal{D}^+_{i,A}\right|^2 - i\,\mathcal{D}^+_{i,A}\mathcal{D}^{+*}_{i,B} + i\,\mathcal{D}^{+*}_{i,A}\mathcal{D}^+_{i,B} + \left|\mathcal{D}^+_{i,B}\right|^2 \\
&= \left|\mathcal{D}^+_{i,A}\right|^2 + \left|\mathcal{D}^+_{i,B}\right|^2 - i\left(\mathcal{D}^+_{i,A}\mathcal{D}^{+*}_{i,B} - \mathcal{D}^{+*}_{i,A}\mathcal{D}^+_{i,B}\right). \quad \textit{Eq. 54}
\end{aligned}$$

This is the celebrated two slits intensity profile (see Eq. 1, Eq. 49, or Young's double-slit experiment in Vol. I) with the third term of the last line being the interference term. (One can show it to be equivalent to the interference term of Eq. 49 sifted by a $-\frac{\pi}{2}$ phase, a proof we omit here because it is not essential to the present discussion.) The negative signature represents an anti-fringe interference pattern.

If one repeats exactly the same calculation, but for Eq. 53, that is, in wavefunction notation for:

$$\Psi^1_{D_2} = \mathcal{D}^+_{\mathrm{i},A} - i\,\mathcal{D}^+_{\mathrm{i},B},$$

then one obtains:

$$|\Psi^1_{D_2}|^2 = |\mathcal{D}^+_{i,A}|^2 + |\mathcal{D}^+_{i,B}|^2 + i\,(\mathcal{D}^+_{i,A}\mathcal{D}^{+*}_{i,B} - \mathcal{D}^{+*}_{i,A}\mathcal{D}^+_{i,B}). \quad \textit{Eq. 55}$$

The positive signature of the interference term represents a fringe interference pattern. A comparison of Eq. 54 and Eq. 55 indicates that what distinguishes the measurements between the two detectors is only the signature in front of the interference term. This might suggest, at first, that we should see a symmetric interference pattern on one detector and an anti-symmetric pattern on the other one.

However, we must repeat the same calculations (this time, we will furnish only the results; check yourself as exercise) by also including the other two possible idler photon states that result after collapse at detector D_0—namely, from the anti-symmetric states $|\Psi_{\mathrm{i}}\rangle^{\mathrm{Asym}}_1$ and $|\Psi_{\mathrm{i}}\rangle^{\mathrm{Asym}}_2$ (Eq. 11b and Eq. 13b).

That of Eq. 11 being:

$$|\Psi_i\rangle^{Asym}_1 = \frac{|-45°_i\rangle_A - |-45°_i\rangle_B}{\sqrt{2}}. \quad \textit{Eq. 56}$$

Therefore, detector D_1 will measure the quantum state:

$$|\Psi_{D_1}\rangle^2 = \frac{-i|-45°_i\rangle_A - |-45°_i\rangle_B}{\sqrt{2}}. \quad \textit{Eq. 57}$$

In wavefunction notation:

$$\Psi^2_{D_1} = -i\,\mathcal{D}^-_{\mathrm{i},A} - \mathcal{D}^-_{\mathrm{i},B},$$

leads to the intensity profile at detector D_1:

$$|\Psi^2_{D_1}|^2 = |\mathcal{D}^-_{i,A}|^2 + |\mathcal{D}^-_{i,B}|^2 + i\,(\mathcal{D}^-_{i,A}\mathcal{D}^{-*}_{i,B} - \mathcal{D}^{-*}_{i,A}\mathcal{D}^-_{i,B}). \quad \textit{Eq. 58}$$

Whereas, detector D_2 will measure the quantum state:

$$|\Psi_{D_2}\rangle^1 = \frac{|-45°_i\rangle_A + i\,|-45°_i\rangle_B}{\sqrt{2}}. \quad \textit{Eq. 59}$$

In wavefunction notation:

$$\Psi^2_{D_2} = \mathcal{D}^-_{\mathrm{i},A} + i\,\mathcal{D}^-_{\mathrm{i},B},$$

leads to the intensity profile at detector D_2:

$$|\Psi_{D_2}^2|^2 = |\mathcal{D}_{i,A}^-|^2 + |\mathcal{D}_{i,B}^-|^2 - i(\mathcal{D}_{i,A}^-\mathcal{D}_{i,B}^{-*} - \mathcal{D}_{i,A}^{-*}\mathcal{D}_{i,B}^-). \quad Eq.\ 60$$

Again, notice the signature difference of the interference term between Eq. 58 and Eq. 60.

Finally, the anti-symmetric part of Eq. 13 being:

$$|\Psi_i\rangle_2^{Asym} = \frac{-|-45°_i\rangle_A + |-45°_i\rangle_B}{\sqrt{2}}, \quad Eq.\ 61$$

then detector D_1 will measure the quantum state:

$$|\Psi_{D_1}\rangle^3 = \frac{i|-45°_i\rangle_A + |-45°_i\rangle_B}{\sqrt{2}}.$$

In wavefunction notation:

$$\Psi_{D_1}^3 = i\,\mathcal{D}_{i,A}^- + \mathcal{D}_{i,B}^-\,,$$

leads to the intensity profile at detector D_1:

$$|\Psi_{D_1}^3|^2 = |\mathcal{D}_{i,A}^-|^2 + |\mathcal{D}_{i,B}^-|^2 + i\,(\mathcal{D}_{i,A}^-\mathcal{D}_{i,B}^{-*} - \mathcal{D}_{i,A}^{-*}\mathcal{D}_{i,B}^-). \quad Eq.\ 62$$

Whereas, detector D_2 will measure the quantum state:

$$|\Psi_{D_2}\rangle^3 = \frac{-|-45°_i\rangle_A - i\,|-45°_i\rangle_B}{\sqrt{2}}.$$

In wavefunction notation:

$$\Psi_{D_2}^3 = -\mathcal{D}_{i,A}^- - i\,\mathcal{D}_{i,B}^-\,,$$

leads to the intensity profile at detector D_2:

$$|\Psi_{D_2}^3|^2 = |\mathcal{D}_{i,A}^-|^2 + |\mathcal{D}_{i,B}^-|^2 - i(\mathcal{D}_{i,A}^-\mathcal{D}_{i,B}^{-*} - \mathcal{D}_{i,A}^{-*}\mathcal{D}_{i,B}^-). \quad Eq.\ 63$$

With the opposite signature difference of the interference term between Eq. 62 and Eq. 63, as expected.

Wrapping it all up first in words, four cases can occur.

When the signal photon collapses onto an anti-symmetric wavefunction of its diagonal basis (anti-fringe – Eq. 10b or Eq. 12b), the idler photon reduces to the symmetric state (Eq. 11a or Eq. 13a; see also Eq. 51), after reflections/transmissions at the mirrors/beamsplitter it is transformed (Eq. 52 or Eq. 53), and, in both cases, 'falls' onto an anti-fringe of detector D_1 (Eq. 54) and a fringe of detector D_2 (Eq. 55).

When the signal photon collapses onto the symmetric wavefunction of its diagonal basis (fringe – Eq. 10a or Eq. 12a), the idler photon reduces to one of two possible anti-symmetric states (Eq. 11b or Eq. 13b; see also Eq. 56 or Eq. 61), after reflections/transmissions at the mirrors/beamsplitter it is transformed (Eq. 57 or Eq. 59) and, in both cases, R_{02} 'falls' onto a fringe of detector D_1 (Eq. 58 or Eq. 62) and an anti-fringe of detector D_2 (Eq. 60 or Eq. 63).

Restating the above in symbols:

$|\Psi_s\rangle_1^{\text{Asym}} \to |\Psi_i\rangle_1^{\text{Sym}} \to$ D_1: anti-fringes; D_2: fringes
$|\Psi_s\rangle_2^{\text{ASym}} \to |\Psi_i\rangle_2^{\text{Sym}} \to$ D_1: anti-fringes; D_2: fringes

$|\Psi_s\rangle_1^{\text{Sym}} \to |\Psi_i\rangle_1^{\text{Asym}} \to$ D_1: fringes; D_2: anti-fringes
$|\Psi_s\rangle_2^{\text{Sym}} \to |\Psi_i\rangle_2^{\text{Asym}} \to$ D_1: fringes; D_2: anti-fringes } R_{01}

where R_{01} and R_{02} indicate the joint detection rates explained in the text.

A IV. Hawking radiation and black hole entropy

The Hawking black body radiation temperature of Eq. 40 is nothing other than the Unruh temperature. One simply plugs into Eq. 39 the gravitational acceleration of a body of mass M at the EH, which, according to Newton's gravitational law, is: $a = GM/R_S^2$, with R_S the Schwarzschild radius of Eq. 38. This delivers the middle term of Eq. 40.

In SI units, one should express masses in kg. Considering that the mass of BHs is of the order of solar masses, that is, $1M_\odot = 2 \times 10^{30}\,kg$, it is more convenient to express Eq. 40 in solar masses. This is readily done if you calculate the quantity $\frac{\hbar c^3}{8\pi\, k_B G} = 1.2275 \times 10^{23}$ (Kelvin times kg in SI units), which implies that it must be rescaled by a factor $1.2275 \times 10^{23} / \, 2 \times 10^{30} = 6.16 \times 10^{-8}$ to finally furnish the right-hand side of Eq. 40.

As to the BH entropy, we proceed as follows. If we insert into the differential expression for the entropy given by Eq. 22 (setting equality) the Unruh temperature of Eq. 39, we obtain:

$$dS = \frac{dQ}{T_{Unruh}} = \frac{Gk_B}{\hbar c^3}\, 8\pi M\, dQ\ .$$

Making use of Einstein's mass-energy equivalence, we can identify the heat dQ with the mass-energy dM falling into the BH and contributing to its entropy increase dS_{BH} as: $dQ = dM \cdot c^2$. Then:

$$dS_{BH} = \frac{Gk_B}{\hbar c}\, 8\pi\, MdM = \frac{Gk_B}{\hbar c}\, d(4\pi M^2),$$

where, in the last step, we made use of the fact that $dM^2 = 2MdM$. Plugging in the mass contained in the BH by means of Eq. 38, namely $M = \frac{R_S\, c^2}{2G}$, one gets:

$$dS_{BH} = \frac{k_B c^3}{\hbar G}\frac{1}{4}\, d(4\pi R_S^2).$$

Considering that, in general, $d(4\pi R^2) = dA$ is the surface differential of a sphere of radius R and area A, and integrating over it the entropy, one finally gets Eq. 41:

$$S_{BH} = \frac{k_B c^3}{\hbar G}\frac{A}{4}.$$

A V. Simple derivation of Planck's scale constants

A simple way to calculate the vacuum quantum fluctuations that are strong enough to form a virtual micro-BH is to begin with the time-energy uncertainty relation (see Vol. I),

$$\Delta E \cdot \Delta t \approx \frac{\hbar}{2}.$$

This is legitimate because, in general, gravitational fields can also be expressed through Einstein's matter-energy equivalence. A volume containing energy in any form is a source of a gravitational field as well. We ask, then, the question: How small must a region of empty space be so that the energy uncertainty is large enough to bring a micro-BH into existence? This is equivalent to asking how short the time interval Δt must be to have an energy uncertainty ΔE that is large enough to contain a matter-energy content that will form a BH. Let us define this small region by a length l_P, or 'Planck length'. Then the short time interval Δt must be that which light needs to travel through this length, that is,

$$t_P = l_P/c.$$

On the other hand, if this length is just that which contains a BH, it must be equal to the Schwarzschild radius of the EH that the energy fluctuation forms. If we identify the mass of the micro-BH, the Planck Mass M_P, in terms of its matter energy equivalence with Planck energy $E_P = M_P c^2$, then, by setting the Planck length equal to the Schwarzschild radius of Eq. 38, one gets

$$l_P = \frac{2GM_P}{c^2} = \frac{2GE_P}{c^4},$$

and from which follows immediately

$$t_P = \frac{l_P}{c} = \frac{2GE_P}{c^5} \quad \text{and} \quad E_P = \frac{l_P c^4}{2G}.$$

Identifying the Planck energy and length, E_P and l_P, with the energy and time uncertainty ΔE *and* Δt of the quantum time-energy relation, one has:

$$E_P\, t_P = \frac{l_P c^4}{2G} \cdot \frac{l_P}{c} \approx \frac{\hbar}{2},$$

From which, finally, follows that (use SI-units in A VI):

$$l_P = \sqrt{G\hbar/c^3} = 1.6 \times 10^{-35} m.$$

Then the other Planck scale constants follow:

$$t_P = \frac{l_P}{c} = \sqrt{G\hbar/c^5} = 0.54 \times 10^{-43} s\,;$$
$$M_P = \frac{E_p}{c^2} = \frac{l_P c^2}{2G} = \sqrt{\hbar c/4G} = 1.1 \times 10^{-8}\text{kg}\,.$$

There are several other, more rigorous derivations of Planck's scale constants (which might also differ slightly for a multiplicative factor from the one given here) but these values are speculative. We don't know what the real physics is when it transitions to a QG regime. These values are supposed to suggest only the order of magnitude where one must expect the known physical laws to break down.

A VI. Physical constant in SI Units

The International System of physical Units (SI Units) is generally used throughout this book. The SI-base units are the meter (m) for length, the kilogram (kg) for mass, the second (s) for time and the Kelvin (K) for temperature. The SI-derived units for energy is the Joule (J) and the Coulomb (C) for the electric charge. Several others exist we however do not need in the present treatise.

Name	Symbol	Value
Speed of light in vacuum 2	c	$299792458\ \frac{m}{s}$
Planck's constant	h	$6.626 \times 10^{-34}\text{Js}$
Newton's gravitational constant	G	$6.674 \times 10^{-11}\frac{\text{m}^3}{kg\,s^2}$
Gravitational acceleration at the Earth's surface	g	$9.81\frac{m}{s^2}$
Boltzmann's constant	k_B	$1.381 \times 10^{-23}\frac{J}{K}$
Avogadro's number	N_A	6.022×10^{23}
Electron (and proton) charge	e	$1.602 \times 10^{-19}\text{C}$
Electron's mass	m_e	$9.109 \times 10^{-31}\text{kg}$
Proton's mass	m_p	$1.672 \times 10^{-27}\text{kg}$

VIII. Acronyms

AdS: anti-de Sitter space
BBO: beta-barium-borate
BEC: Bose-Einstein condensate
BH: black hole
BM: Bohmian mechanics
BW: band-width
BSM: Bell-state measurement
CFT: conformal field theories
CM: classical mechanics
CMB: cosmic microwave background
CP: classical physics
DCQE: delayed choice quantum eraser
EH: event horizon
EM: electromagnetic
EPR: Einstein-Podolsky-Rosen
EW: electroweak
FTL: faster than light
GR: general relativity
QB: quantum biology
QC: quantum computer
QCD: quantum chromo-dynamics
QED: quantum electrodynamics
QFT: quantum field theory
QG: quantum gravity
QLG: quantum logic gate
QM: quantum mechanics
QP: quantum physics
QT: quantum theory
MSG: modified Stern-Gerlach
MWI: many world interpretation
MZI: Mach-Zehnder interferometer
SEW: Scully, Englert and Walther
SG: Stern-Gerlach
SM: standard model (of particle physics)
SPDC: spontaneous parametric down-conversion
SR: special relativity
ST: string theory
TI: transactional interpretation of QM
WMAP: Wilkinson Microwave Anisotropy Probe
ZWM: Zou, Wang, Mandl (experiment)

IX. Acknowledgements

I'm particularly grateful to Mark Kelleher, a scientist working in an Antarctic research center, who reviewed this document by making a careful proofread. I would also like to offer my special thanks to Tonya Blust; she is an exceptional proofreader (https://www.fiverr.com/bluston) whom I surely recommend. Without their help, this book would have sounded ridiculous!

X. Bibliography

[1] L. J. W. a. L. M. X. Y. Zou, "Induced coherence and indistinguishability in optical interference," *Physical Review Letters,* vol. 67, p. 318, 1991.

[2] D. Ellerman, "Why delayed choice experiments do Not imply retrocausality," *Quantum Stud.: Math. Found.,* vol. 2, p. 183, 2015.

[3] J. Fankhauser, "Taming the delayed choice quantum eraser".

[4] B. Gaasbeeck, "Demystifying the delayed choice experiments," 2010.

[5] R. E. Kastner, "The delayed choice quantum eraser neither erases nor delays," *Foundations of Physics,* vol. 49, no. 7, p. 717, 2019.

[6] Y. Y. F. e. al., "Quantum superposition of molecules beyond 25 kDa," *Nature Physics,* 23 September 2019.

[7] B.-G. E. &. H. W. Marian O. Scully, "Quantum optical tests of complementarity," *Nature,* vol. 351, p. 111–116, 1991.

[8] M. O. T. C. S. P. a. C. H. M. S. P. Walborn, "Double-slit quantum eraser," *Phys. Rev. A,* vol. A, no. 65, p. 033818, 2002.

[9] Y.-H. Kim, R. Yu, S. P. Kulik, Y. H. Shih and M. Scully, "A Delayed "Choice" Quantum Eraser," *Physical Review Letters,* vol. 84, no. 1, pp. 1-5, 2000.

[10] R. A. Robert Brady, "Why bouncing droplets are a pretty good model of quantum mechanics," 2014.

[11] N. Wolchover, "Famous Experiment Dooms Alternative to Quantum Weirdness," 11 October 2018. [Online]. Available: https://www.quantamagazine.org/famous-experiment-dooms-pilot-wave-alternative-to-quantum-weirdness-20181011/.

[12] D. Bohm, Wholeness and the Implicate order, Roudledge, 1980.

[13] R. E. Kastner, "Is there really "retrocausation" in time-symmetric approaches to quantum mechanics?," *AIP Conference Proceedings,* vol. 1841:1, 2017.

[14] R. E. Kastner, The Transactional Interpretation of Quantum Mechanics: The Reality of Possibility, Cambridge University Press, 2012.

[15] B. S. Finn, “Laplace and the speed of sound,” *Isis,* vol. 55, no. 1, pp. 7-19, 1964.

[16] [Online]. Available: https://gfycat.com/gifs/detail/FickleSorrowfulCommabutterfly.

[17] S. Phelps, “GeoGebra,” [Online]. Available: https://alpha.geogebra.org:446/u/stevephelps.

[18] P. Woit, Not Even Wrong: The Failure of String Theory and the Search for Unity in Physical Law, Basic Books, 2007.

[19] S. Hosselfelder, Lost in Math: How Beauty Leads Physics Astray, Basic Books, 2018.

[20] M. Masi, Free Progress Education, https://www.amazon.com/gp/product/B07SVFZD3Z/.

[21] X.-S. M. e. al., "Quantum teleportation over 143 kilometres using active feed-forward," *Nature,* vol. 489, p. 269–273, 2012.

[22] S. Bushwick, "New Encryption System Protects Data from Quantum Computers," 8 October 2019.

[23] M. Masi, "Free Progress Education", 2017, CreateSpace Independent Publishing Platform - Online: https://www.amazon.com/dp/1539673081.

[24] Venkataraman, Bose And His Statistics, Universities Press, 1992, p. 14.

[25] "NIST/JILA/CU-Boulder - NIST Image," [Online].

[26] T. M. D. K. a. K. Andrews, "Observation of Interference Between Two Bose Condensates," *Science,* Vols. 637-41, pp. 275-5300, 1997.

[27] N. S. T. -. W. 101087, "WMAP webpage," Nasa, 16 04 2010. [Online]. Available: https://wmap.gsfc.nasa.gov/universe/bb_tests_ele.html.

[28] P. J. S. A. L. Anna Ijjas, "Cosmic Inflation Theory Faces Challenges," *Scientific American,* January 2017.

[29] "A Cosmic Controversy," *Scientific American,* February 2017.

[30] J. A.-K. a. J. McFadden, Life on the Edge: The Coming of Age of Quantum Biology, Bantam Press, 2014.

[31] H. Stapp, Quantum Theory and Free Will: How Mental Intentions Translate into Bodily Actions, Springer, 2017.

[32] R. P. Stuart Hameroff, "Consciousness in the universe: A review of the 'Orch OR' theory'," *Phys. Life Rev.,* vol. 11, pp. 39-78, 2014.

[33] S. K. e. al., "Possible existence of optical communication channels in the brain," *Scientific Reports,* vol. 6, p. 36508, 2016.

[34] C. Simon, "Can quantum physics help solve the hard problem of consciousness? A hypothesis based on entangled spins and photons.".

[35] "MIT's replication of Yves Couder's wave-guided particle effect," [Online]. Available: https://www.youtube.com/watch?v=YF5iHQMjcsM.

XI. About the author

Marco Masi was born in 1965 and attended the German School of Milan, Italy. He graduated in physics at the university of Padua, and later obtained a Ph.D. in physics at the university of Trento. He worked as a postdoc researcher in universities in Italy, France, and more recently in Germany, where he worked also as a school teacher for three years. After he had authored some scientific papers (www.researchgate.net/profile/Marco_Masi2), his interests veered towards new forms of individual learning and a new concept of free progress education originated from his activity both as a tutor in several universities and as a high school teacher, but especially from his direct, lived experience of what education should **not** be. From this originated also his desire to write this book on QM which tries to close a gap between the too high-level university textbooks and a too low level popular science approach that is so typical of modern hyped media. He is also interested in metaphysical and philosophical ruminations and loves walking in the woods, loves animals and would never kill a cat to carry out Schrödinger's experiment.

XII. Index